



Mikko Iivarinen

# Vuodepaikkojen määrän ennakointi sairaalaympäristössä

## Kirjallisuuskatsaus

Metropolia Ammattikorkeakoulu

Insinööri (AMK)

Tieto- ja viestintätekniikka

Insinöörityö

31.10.2021

## Tiivistelmä

Tekijä: Mikko Iivarinen  
Otsikko: Vuodepaikkojen määrän ennakointi sairaalaympäristössä  
Sivumäärä: 41 sivua + 4 liitettä  
Aika: 31.10.2021

Tutkinto: Insinööri (AMK)  
Tutkinto-ohjelma: Tieto- ja viestintätekniikka  
Ammatillinen pääaine: Hyvinvointi- ja terveysteknologia  
Ohjaajat: Lehtori Sakari Lukkarinen  
Laatupäällikkö Mari Järvisalo

---

Insinööriyössä tehtiin kirjallisuuskatsaus olemassa oleviin ennakoiviin malleihin ja niissä käytettäviin algoritmeihin ja parametreihin. Tilaajan tavoitteena työssä oli teoriapohjan luominen ennakoiviin malleihin sekä se, millaista tietoa Espoon sairaalan tulisi kerätä mallin toteuttamiseksi.

Työssä tutkittiin lean-ajattelun perusteita, kirjallisuuskatsaustyyppisiä ja tarkemmin scoping-katsausta, jonka periaatteiden mukaisesti kirjallisuuskatsaus toteutettiin. Järjestelmällisen tiedonhaun aikana aineistoa haettiin kuudesta tunnetusta tietokannasta ja viitteet tallennettiin ProQuest RefWorks -ohjelmaa käyttäen. Tiedonhaun vaiheet ja aineiston analysointi taltioitiin Microsoft Excel -taulukkolaskentaohjelmalla.

Aineisto analysoitiin laadullisen sisällönanalyysin koodausmenetelmää käyttäen Excel-taulukkoon. Tärkeimmät tutkimuksista kerätyt tiedot liittyivät käytetyn mallin tyyppiin, mihin algoritmiin malli perustui sekä millaisia parametreja algoritmi hyödynsi. Tutkimuksien otantojen ympäristöstä riippuen myös ennakoivien mallien tarkkuudet kirjattiin ylös.

Tarkimmat koneoppimismenetelmiin perustuvat ennakoivat mallit saavuttivat omissa ympäristöissään yli 90 %:n tarkkuuden hoitajaksojen pituuksien ja kotiuttamisajan kohdan arvioinnissa. Osa malleista pystyi ennakoimaan jopa järjestyksen, jossa potilaat kotiutuivat osastolta. Tutkimusten analysoinnin aikana huomattiin, että mallien tarkkuudet eivät ole lineaarisessa suhteessa käytettyjen parametrien määrään.

Insinööriyön tuloksena syntyi kattava tietopaketti ennakoivista malleista niin tämän raportin kuin tulostaulukonkin muodossa. Espoon sairaalaa konsultoitiin työn tuloksista, niiden perusteella sairaala voi aloittaa omaan ympäristöönsä sopivan ennakoivan mallin kehittämisen.

Avainsanat: Espoon sairaala, lean, ennakoiva malli, scoping-katsaus, vuodepaikka

## Abstract

Author: Mikko Iivarinen  
Title: Forecasting Hospital Bed Occupancy – Scoping Review  
Number of Pages: 41 pages + 4 appendices  
Date: 31 October 2021

Degree: Bachelor of Engineering  
Degree Programme: Information and Communication Technology  
Professional Major: Health Technology  
Supervisors: Sakari Lukkarinen, Senior Lecturer  
Mari Järvisalo, Quality Manager

---

The purpose of the study was to do a literature review on existing predictive models used in health care and study different algorithms and parameters used in those models. The objective of the study was to create a theoretical basis about predictive models and what kind of information would Espoo Hospital need to collect to implement a predictive model suitable for their needs.

This study examined lean-principles, different types of literature reviews and more precisely scoping review, which was the chosen literature review method. Systematic search was used to collect material from six different known databases and the references were saved using ProQuest RefWorks software. Phases of the systematic search and analysing the collected material were done using a Microsoft Excel spreadsheet.

The material was analysed using a qualitative content analysis technique called coding and the categories were recorded in an Excel spreadsheet. The most important pieces of information from the different studies were the types of predictive models, the algorithms the models used, and the parameters collected and used. Accuracies of the predictive models were also recorded if the study setting made it possible.

The most accurate predictive models based on machine learning techniques achieved over an 90 % accuracy when forecasting length of stay and discharge times in optimal study settings. Some of the models were even able to forecast the order in which patients were discharged from their departments. It was also discovered during the study that the accuracies of the predictive models were not linearly relative to the used parameters.

The outcome was a broad study about predictive models including this report and the resulting table of analysed studies. Espoo Hospital was consulted about the results and based on the information the hospital may start developing a predictive model suitable for their needs.

Keywords: Espoo Hospital, lean, predictive model, scoping review, hospital beds

# Sisällys

## Lyhenteet

1	Johdanto	1
2	Espoon sairaala	2
2.1	Kuntoutus	3
2.2	Potilasmäärien ennakointi	4
3	Lean-ajattelu	4
3.1	Historia ja konseptit	4
3.2	Hukka	6
3.3	Moderni lean-ajattelu	7
4	Opinnäytetyön tarkoitus, tavoite ja tutkimuskysymys	8
5	Kirjallisuuskatsaus	9
5.1	Kirjallisuuskatsauksen tyypit	9
5.1.1	Kuvailevat kirjallisuuskatsaukset	9
5.1.2	Systemaattiset kirjallisuuskatsaukset	10
5.2	Scoping-katsaus	11
5.3	Tiedonhaun teoria	13
5.4	Haku- ja valintakriteerit	15
5.5	Aineiston etsiminen	16
5.6	Valitut aineistot	17
5.7	Aineiston arviointi	26
5.8	Aineiston analysointi	26
6	Tulokset	29
6.1	Aineistosta saadut tiedot	29
6.2	Konsultointi	31
7	Yhteenveto	32
7.1	Tulosten tarkastelu ja pohdinta	32
7.2	Luotettavuus ja eettisyys	34
7.3	Mitä opin	34

7.4 Jatkotutkimus ja ehdotukset	35
Lähteet	36
Liite 1: Kokotekstivalinnan artikkelit	
Liite 2: Tulostaulukko analysoiduista artikkeleista	
Liite 3: Kabir, S. & Farrokhvar L. tutkimuksen parametrit	
Liite 4: Kabir, S. & Farrokhvar L. tutkimuksen tarkkuudet	

## Lyhenteet

5S	Sort, set in order, shine, standardized control, sustain. Selvitä, sijoita, siisti, standardisoi, säilytä.
ANN	Artificial neural network. Keinotekoinen neuroverkko, eräs koneoppimismenetelmä.
AP	Average precision. Keskimääräinen tarkkuus, eräs koneoppimisessa käytetty mittari mallin tarkkuudesta.
ASA	American Society of Anesthesiologists. Viittaa yleisesti ASA-luokitukseen eli fyysisen tilan luokitusjärjestelmään.
AUROC	Area under receiver operating characteristic. ROC-käyrän alapuolelle jäävä pinta-ala.
CT	Classification tree. Päätöksentekopuu, eräs koneoppimismenetelmä ja ennakoiva mallintamistapa.
EWS	Early warning score. Peruselintoimintoja mittaava pisteytys.
FL	Fuzzy Logic. Sumea logiikka, eräs koneoppimismenetelmä.
GCS	Glasgow Coma Scale. Glasgow'n kooma-asteikko. Käytetään lääketieteessä henkilön tajunnantason arviointiin.
HAN	Hierarchical Attention Network. Luonnollisen kielen automaattisessa prosessoinnissa käytetty koneoppimismenetelmä.
HTK	Hyvä tieteellinen käytäntö.
ICD-9	International Classification of Diseases, Ninth revision. Vanha lääketieteellinen tautiluokitus.

JIT	Just-in-Time. Yksi lean-ajattelun konsepteista. Tarkoittaa tiettyä määrää täydellistä tuotetta tietyssä paikassa.
LoS	Length of stay. Hoitojakson pituus.
MAPE	Mean absolute percentage error. Keskimääräinen absoluuttinen prosentuaalinen virhe tai poikkeama. Yleinen ennustetarkkuuden mittari.
mRS	Modified Rankin Score. Omatoimisuutta mittaava asteikko.
MSE	Mean squared error. Keskineliövirhe, ennakoinnissa käytetty virhemittari.
NLP	Natural language processing. Luonnollisen kielen käsittelyjärjestelmä.
NSQIP	National Surgical Quality Improvement Program. Yhdysvaltalainen leikkausten laadun parantamiseen keskittyvä järjestelmä.
PICO	Patient, Intervention, Comparison, Outcome. Potilas, interventio, vertailuinterventio, hoidon tulos. Terveystieteiden hakutermit.
RF	Random Forest. Satunnaismetsämalli, eräs koneoppimismenetelmä.
ROC	Receiver operating characteristic curve. Binääriluokittelussa käytettävä kuvaaja, kertoo mallin tarkkuudesta.
RRF	Regression random forest. Satunnaismetsämalli jatkettuna regressioanalyysillä, yhdistetty koneoppimismenetelmä.
SALSA	Search, Appraisal, Synthesis, Analysis. Haku, arviointi, synteesi, analyysi. Kirjallisuuskatsauksen normaalit osiot.

SIR	Susceptible, Infected, Recovered. Sairastumiselle alttiit, tartuttavat, toipuneet. SIR-mallia käytetään epidemian tai tartuntataudin matemaattiseen ennustamiseen.
SOFA	Sequential Organ Failure Assessment score. Elimien toimintahäiriötä mittaava pistejärjestelmä.
TENK	Tutkimuseettinen neuvottelukunta.
TPS	Toyota Production System. Toyotan tuotannon ohjausjärjestelmä.



# 1 Johdanto

Terveydenhuolto on jatkuvasti haasteellisessa muutoksessa. Syitä on monia, ja pahimmillaan ne voimistavat toisiaan. Suomen ja Euroopan väestörakenne vanhenee [1], mikä pahentaa jo olemassa olevaa hoitajapulaa sekä kasvattaa terveydenhuollon raskautta ja kuluja [2]. Lisäksi ulkoiset tekijät, kuten COVID-19-pandemia, aiheuttavat suuria ongelmia. Talvella 2021 uutisoitiin potilaiden hoidon ja leikkausten siirtymisestä myöhempään koronapotilaiden määrän kasvun vuoksi [3].

Väestörakenteen muutoksen tuomat ongelmat ovat olleet tiedossa jo vuosia, ja se on tuonut paljon huomiota ja kiinnostusta ennakoivia malleja kohtaan. Terveydenhuollossa ennakkoinnilla pyritään potilasmäärien ja hoitajaksojen pituuk-sien ennustamiseen ruuhkien välttämiseksi sekä kulujen ja resurssien tasapainottamiseksi. Tämä antaa myös mahdollisuuden tehokkaaseen työvuorosuunnitteluun. [4.]

Opinnäytetyön tilaajana on Espoon sairaala. Työn tarkoituksena on tehdä kirjallisuuskatsaus olemassa olevaan teoretiseen ennakoivista malleista. Tavoitteena on tunnistaa potilasmäärien ennakkointiin ja hoitajaksojen pituuteen vaikuttavia historiallisia ja reaaliaikaisia parametrejä sekä erilaisia ennakoivien mallien tyyppisiä. Kirjallisuuskatsauksen tavoitteena on tuottaa tietoa Espoon sairaalalle ennakoivien mallien mahdollisuuksista sekä konsultoida Espoon sairaalaa siitä, millaista potilastietoa tulisi mahdollisesti kerätä ja käyttää.

Opinnäytetyön taustalla on Lean-ajattelu, jota Espoon sairaala hyödyntää johtamisfilosofiassaan. Työ on osa Lean-ajattelun jatkuvan kehittämisen periaatetta (motion study) [5, s. 2; 84], jossa teknologiaa yritetään hyödyntää mahdollisimman monipuolisesti samalla hukkaa vähentäen. Työssä käydään läpi Lean-ajattelun perusteita ja historiaa, kirjallisuuskatsaustyyppisiä, tiedonhaun perusteita ja tietokantoja sekä valittujen tutkimusten analysointia. Lopuksi nostetaan esiin tärkeimpiä mukana olleita tutkimuksia sekä pohditaan nykytilannetta ja mahdollista jatkotutkimusta.

## 2 Espoon sairaala

Espoon sairaala (kuva 1) valmistui Jorvin sairaalan viereen vuonna 2016, se aloitti toimintansa 7. maaliskuuta 2017 [6, 7]. Espoon sairaalassa on kuusi osastoa: päivystysosasto, haavanhoito- ja psykogeriatrinen kuntoutusosasto, ortopedinen kuntoutusosasto, neurologinen kuntoutusosasto, infektio-osasto sekä Villa Glims, jossa järjestetään palliatiivista hoitoa ja saattohoitoa. [8.]

Espoon sairaala on erikoistunut ikääntyneiden hoitoon ja kuntoutukseen sekä kotona asumisen tukemiseen. Sairaalassa hoidetaan Espoolaisten lisäksi Kau-  
niaisten ja Kirkkonummen ikääntyneitä. Työntekijöitä on yli 400 ja potilashuoneita 247. [9, 10.]



Kuva 1. Espoon sairaala Jorvin sairaalakampuksella, Karvasmäentie 6 [13].

## 2.1 Kuntoutus

Espoon sairaalan osastot ovat monipuolisia kuntoutusosastoja, ja toiminnan keskiössä on omatoimisuuden tukeminen [11].

Espoon sairaalassa hoidon periaatteena on potilaan omatoimisuuden tukeminen kuntouttavan työotteiden avulla. Kuntouttavassa työotteessa potilasta kannustetaan tekemään itse mahdollisimman paljon niitä asioita, joita hän pystyy tekemään. Potilas saa apua niihin asioihin, joita hän ei pysty tekemään. [11.]

Jokaiselle potilaalle tehdään yksilöllinen hoito- ja kuntoutussuunnitelma sekä asetetaan konkreettiset tavoitteet. Toimintakyky pysyy yllä, kun potilaalle annetaan mahdollisuus määrittää oman aktiivisuutensa taso ja kuntoutus aloitetaan mahdollisimman nopeasti. Kuntoutusta ja terapiaa järjestetään niin yksilö- kuin ryhmäterapiainakin [10]. Sairaalassa on ruokailutilat, yhteisiä ajanvietto- ja kuntoutustiloja sekä ohjattua toimintaa. Tiloista löytyy erilaisia kuntoilulaitteita (kuva 2), aktivoivia pelejä ja pelikonsoleita sekä katettu sisäpiha ulkoiluun. [11, 12.]



Kuva 2. Kuntolaitteita Espoon sairaalassa [13].

## 2.2 Potilasmäärien ennakointi

Tällä hetkellä Espoon sairaala saa etukäteen tiedon muista yksiköistä saapuvista potilaista joko suoraan HUSin Jorvin sairaalan yhteispäivystyksestä tai hoivonvarausjonosta Navitas-tietojärjestelmästä. Historiallista tietoa olisi saatavilla, mutta sen hakeminen eri lähteistä reaaliaikaisesti on vaikeaa ja tieto päivittyy hitaasti. Lisäksi ammattilaiset arvioivat kokemuksen perusteella esimerkiksi tiettyjen vuodenaikojen kiireisyyttä. Nopeinta reagointia Espoon sairaalalta vaatii HUSin Jorvin sairaalan yhteispäivystyksen potilasvirtoihin vastaaminen. [13.]

Vaikka jokaiselle Espoon sairaalaan tulevalle potilaalle tehdään henkilökohtainen hoitosuunnitelma, on hoitajaksojen pituuksien arviointi haastavaa. Tämä johtuu siitä, että ihmisten kuntoutuminen ja sen nopeus ovat yksilöllisiä. Lisäksi kuntoutumisen arviointi on vaikeaa eikä siinä käytettävät mittarit toimi kaikilla. [12.] Hoitajaksojen pituus ja vaihtelu tekevät vapaiden vuodepaikkojen laskemisesta ja ennustamisesta vaikeaa.

Espoon sairaalassa toivotaan uusia työkaluja potilasmäärien ennakointiin. Tilailijan toive rajaukselle opinnäytetyössä olisi löytää teoriapohja sille, miten voitaisiin ennustaa potilasmäärää kahden päivän päähän.

Ensimmäinen askel ennakoinnissa tulisi olla hoitajaksojen pituuksien tarkempi arviointi. Liian lyhyeksi arvioitu hoitajakson pituus aiheuttaa jatkuvia muutoksia, eikä haluttuun ennakointiin päästä. Liian pitkäksi arvioidut hoitajaksojen pituudet vääristäisivät tilastoa aiheuttaen jatkuvia virheitä ennakoinnin tarkkuudessa, mikä veisi tilastolta luotettavuuden. Jokaiselle potilaalle tulisi arvioida hoitajakson pituus, jotta voitaisiin laskea ennuste vapaista paikoista osastoilla. [13, 14.]

## 3 Lean-ajattelu

### 3.1 Historia ja konseptit

Lean-ajattelun historia on lähtöisin 1950-luvun Japanista, kun Toyotan johto halusi ymmärtää, miten Fordin autotehtaat pystyivät tuottamaan tuhansia autoja

päivässä, kun Toyotalla tuotettiin vähemmän vuodessa. Lean-ajattelun ytimessä on hukana ja arvoa tuottamattoman työn poistaminen [15]. Lean voidaan jakaa kolmeen konseptiin, joista ensimmäinen on japanilainen kaizen, eli jatkuva kehitys (käännös yleisesti: Muutos parempaa kohti). Yleensä kaizeniksi kutsutaan lyhyttä projektia, jossa eri virkoja hoitavat ihmiset koostavat työryhmän, joka paneutuu yhteen prosessiin ja pyrkii kehittämään sitä. [16.]

Jidoka on toinen lean-konsepteista ja tarkoittaa automaatiota, jolla on käyttäjä tai valvoja (eng. Automation with a human touch). Jidokan mukaan paras automaatio on helppo automaatio, ja tarkoittaa käytännössä jokaisen tuotteen tai vaiheen automaattista tarkistamista niin koneellisesti kuin manuaalisestikin. Ongelmatuotteita tai -vaiheita ei pelkästään poisteta, vaan koko systeemi pysäytetään, kunnes ongelmien aiheuttaja on löydetty ja korjattu. Tätä voidaan hyödyntää ennakoivan mallin rakentamisessa niin, että tietojen syöttö on mahdollisimman helppoa, nopeaa ja pitkälle automatisoitua, kun käytetään esimerkiksi luonnollisen kielen prosessointia (natural language processing, NLP). [17; 5, s. 177, 181.]

JIT (Just-in-Time) on alkuperäisistä lean-konsepteista kolmas ja tiivistettynä tarkoittaa tiettyä määrää, tiettyyn paikkaan, tiettyä aikana täydellisessä kunnossa. JIT on hyödyllisimmillään teollisessa ympäristössä ja tehtaissa, missä vaiheita on paljon tai volyymit isoja. [5, s. 5.]

Toyotan johto yhdisti kaikki kolme lean-ajattelun konseptia: kaizenin, Jidokan ja JIT:n, ja niistä syntyi Toyotan tuotannon ohjausjärjestelmä (Toyota Production System, TPS). Konseptien perusteella käyttöympäristöön suunnitellaan ensin ideaali ratkaisu ja sitten luodaan käytännöllinen ratkaisu. [5, s. 9, 31.]

Lean-ajattelu ja sen periaatteet voidaan käytännössä jakaa viiteen askeleeseen (kuva 3.) [18]:

- Etsitään arvoa asiakkaan näkökulmasta.
- Identifioidaan arvoa tuottavat vaiheet ja poistetaan turhat vaiheet.
- Arvoa tuottavien vaiheiden tiivistäminen ja virtauksen parantaminen.

- Varmistetaan, että arvon virtaus asiakkaalle pysyy ja etsitään seuraava parannettava vaihe.
- Toistetaan, kunnes saavutetaan lähes täydellinen tuotteiden tai palveluiden virtaus ja kaikki hukka on poistettu.



Kuva 3. Lean-ajattelun viiden askeleen periaate [18].

### 3.2 Hukka

Lean-ajattelussa hukalla tarkoitetaan erilaisia turhia tai tuhlaavia asioita, materiaaleja ja työvaiheita. Hukan alkuperäinen määritelmä oli ”kaikki mikä ei ole ehdottoman tärkeää” [5 s. 7]. Erilaisia hukan muotoja oli alun perin seitsemän [5, s. 7], mutta nykyään voidaan todeta kahdeksan:

- ylituotanto (overproduction)
- ylivarastointi (inventory)
- turhat siirrot ja kuljetukset (transportation)
- vialliset tuotteet ja niiden korjaaminen (defects)
- turhat prosessit ja työvaiheet (processes)
- tarpeettomat ja hyödyttömät tuotannon tai yrityksen osat (operations)
- turha työkoneiden seisotus ja työntekijöiden odottelu (inactivities)

- kehityspotentiaalin hukkaaminen ja tiedon hyödyntämättä jättäminen [15].

Hukan muodoista ylivarastointia on yleisesti pidetty suurimpana vaikuttajana tai hukkana. On helppoa aliarvioida varastoinnin kustannukset: kuljetus varastoon, varaston rakentaminen, vuokraus tai maan hinta, varastotyöntekijöiden managerointi ja varastotyökoneiden käyttökustannukset ja huollot. [5 s. 7.] Hukan eri muodot on muutettu myös terveydenhuoltoalalle jatkuvan kehityksen ja kasvun vuoksi [19]:

- Odotutetaan potilaita ja työntekijöitä turhaan.
- Minimoidaan varastot sisältäen potilaat, tarvikkeet ja datan määrä.
- Järjestelmävirheet, hoitovirheet ja väärät diagnoosit ovat kaikki virheitä ja ajan tuhlausta.
- Liikuttelu: tulee vähentää potilaiden, tarvikkeiden ja henkilöstön liikumista potilasvirran parantamiseksi.
- Liikkumisen vähentäminen lisää työaika ja parantaa turvallisuutta, kun työntekijöiden ei tarvitse kurotella välineitä tai kävellä käytävää edestakaisin.
- Ylituotantoa voi olla myös terveydenhuollossa. Tähän lukeutuu esimerkiksi ylimääräisten lääkkeiden jakelu, testien tai lomakkeiden turha tai uusiutuva täyttäminen, sekä liian pitkät hoitajaksot.
- Prosessien turha uusiutuminen on päällekkäistä ylituotannon kanssa. Edellisten lisäksi myös datan syöttäminen useampiin järjestelmiin on turha prosessi, joka voidaan poistaa.
- Miten muut hukan muodot vaikuttavat työntekijöihin ja heidän ajankäyttöön. Kaikki mikä on pois potilaiden hoitamisesta, on terveydenhuollon näkökulmasta turhaa.

### 3.3 Moderni lean-ajattelu

5S-malli (sort, set in order, shine, standardized control, sustain) [5, s. 148] on ensimmäinen osa JIT:iä. 5S-mallia voidaan käyttää myös tietokonepainotteisessa työympäristössä. Työpistettä ja päätelaitetta tulee puhdistaa ja tyhjentää, jotta se pysyy nopeana ja vähentää turhaa odottelua. Tiedostot ja kansiot tulisi tehdä siisteihin ”puihin”, jotta varmuuskopio tarvitsisi tehdä vain tiedostoista, eikä koko järjestelmästä, mikä vähentää hukkaa ja säästää tilaa. [5, s. 156.]

Terveysthuollon näkökulmasta 5S-malli voidaan suomentaa muotoon ”selvitä, sijoita, siisti, standardisoi, säilytä” [16]. 5S-mallia näkee monessa suomalaisessa yrityksessä, joissa ei lean-ajattelua muuten käytettäisi. Se on hyvä muistisääntö työpisteen siistinä pitämiseen ja laadun ylläpitämiseen.

Ennakoivan mallin rakentamisessa voidaan käyttää myös kahta visuaalista strategiaa: Red-Tagging Strategy ja Sign Strategy, eli punakorttistrategia ja visuaalinen huomiointi. Punakorttistrategia tarkoittaa varoitusten eli ”keltaisten korttien” poistamista, jolloin ei päädytä ikuisen varoitusten kiertoon. Tämä poistaisi ennakoivalta mallilta kaiken hyödyn, kun kukaan ei kiinnittäisi varoituksiin huomiota. Visuaalisen huomioinnin tarkoituksena on käyttää kuvia, logoja ja värejä huomion kiinnittämiseksi ja asioiden löytämiseksi aikaa säästäen. Koska opinäytetyössä luodaan vasta teoriapohjaa ennakoivalle mallille, voidaan visuaalisia asioita käsitellä vasta varsinaisen mallin luonnin yhteydessä. [5 s. 158.]

Viimeisen parin vuosikymmenen aikana on tullut käyttöön hybridimalli Lean Six Sigma. Hybridimallin tarkoituksena on yhdistää kaksi erilaista lähestymistapaa tuotannon ja toiminnan optimointiin. Lean keskittyy hukun poistamiseen, kun taas Six Sigma vaihtelun ja virheiden vähentämiseen tieteellisin ja tilastollisin keinoin. [19.]

Six Sigman tarkoituksena on systemaattinen parantaminen tilastojen perusteella. Vaihtelun ja hukun vähentäminen vähentää virheitä. Alun perin tarkoituksena ja tilastollisena tavoitteena, mistä nimikin on lähtöisin, on saada virheet alle 3,4 ppm eli 3,4 virhettä miljoonassa. [20, s. 5.]

#### **4 Opinnäytetyön tarkoitus, tavoite ja tutkimuskysymys**

Opinnäytetyön tarkoituksena on tehdä scoping-katsaustyyppinen kirjallisuuskatsaus Lean-ajattelun pohjalta ennakoiviin malleihin sairaalaympäristössä potilasvirran ylläpitämiseksi sekä hoitajaksojen pituuksien arvioimiseksi. Tavoitteena työssä on löytää tarvittavat tilastolliset parametrit ja tiedot ennakoivan mallin luomiseksi ja koostaa niistä teoriapohja Espoon sairaalalle. Tilaajan toiveena on



ennakointi kahden päivän päähän. Tutkimuskysymys työssä on: ”Millaista tietoa sairaalan tulisi kerätä ja analysoida ennakoivan mallin luomiseksi ja käyttämiseksi?”

Tiedonhaussa ja tutkimuksien valinnassa keskitytään koneoppiviin malleihin, sillä niiden kehitys ja käyttö on lähivuosina kasvanut huomattavasti [14]. Sairaalat keräävät isoja määriä elektronista potilasdataa jatkuvasti [21]. Monet tutkijat yrittävät luoda tarkkaan ennustavan mallin, joka on nopea ottaa käyttöön.

## 5 Kirjallisuuskatsaus

### 5.1 Kirjallisuuskatsauksen tyypit

Kirjallisuuskatsaukset voidaan luokitella kolmeen eri päätyyppiin: kuvaileviin ja systemaattisiin katsauksiin sekä meta-analyyseyihin. Jokaisella päätyypillä on useampia alatyyppejä, ja uusia luodaan jatkuvasti. [22; 23, s. 8.]

Koska kirjallisuuskatsaus on tutkimustyyppi, tulee sen olla toistettavissa [23, s. 7]. Kirjallisuuskatsauksen tärkein tehtävä on kehittää ja arvioida teoriaa sekä tunnistaa ristiriitoja ja ongelmia. Kirjallisuuskatsauksissa on yleensä neljä tyyppilistä osaa: haku, arviointi, synteesi ja analyysi, jotka muodostavat englanninkielisen muistisäännön SALSA (Search, Appraisal, Synthesis, Analysis) [23, s. 8]. Erityyppisissä kirjallisuuskatsauksissa osien painoarvot ja laajuudet kuitenkin vaihtelevat.

#### 5.1.1 Kuvailevat kirjallisuuskatsaukset

Kuvailevien katsausten tarkoituksena on kuvata aiempaa tutkimusta, syvyyttä, laajuutta ja määrää. Näistä nimensä mukaisesti yleisimmät ovat narratiivinen (narrative literature review) ja perinteinen (traditional literature review) kirjallisuuskatsaus, jotka tarkastelevat ja yhteen vetävät tiettyyn aiheeseen liittyviä tutkimuksia. Luotettavuus tällaisilla katsauksilla vaihtelee, sillä ne eivät arvioi tutkimusten tai niiden lähteiden luotettavuutta. [23, s. 9.]

Kriittinen (critical review) ja kartoittava (mapping review) katsaus ovat eri alatyyppejä, mutta niiden pääasiallinen tehtävä on sama: löytää uuden tutkimuksen tarpeita. Kriittinen katsaus keskittyy tarkemmin tiettyyn aiheeseen ja pyrkii tuomaan esille epäkohtia ja ristiriitaisuuksia. Kartoittava katsaus taas on laajempi, eikä sen tarkoituksena ole tuottaa yhteenvetoa tai varsinaisia tuloksia, vaan kategorisoida tietoa myöhemmälle tutkimukselle. [23, s. 10.]

Scoping-katsaus (scoping review) tehdään yleensä tutkimuksen laajuuden, luonteen ja määrän selvittämiseksi ja kuvaamiseksi. Tarkoituksena ei ole analysoida yksittäisiä tutkimuksia, vaan tuottaa laajempi yhteenveto sekä selvittää tarvetta ja mahdollisuutta systemaattiselle katsaukselle. Scoping-katsaus on myös kriittisempi kuin perinteiset kuvailevat katsaukset, ja aineiston etsimisen järjestelmällisyys on samalla tasolla systemaattisen katsauksen kanssa. [23, s. 11.]

Nopea katsaus (rapid review) on tiettyyn ja tarkoin määriteltyyn aiheeseen, usein käytännönongelmaan, liittyvä katsaus. Sitä käytetään usein strategisessa ja poliittisessa päätöksenteossa, ja siinä voidaan hyödyntää sekä määrällistä että laadullista tutkimusta kontekstista riippuen. [23, s. 12.]

Yleiskatsaus (overview) analysoi ja tiivistää olemassa olevaa tutkimustietoa tietyistä aiheista olematta kovin systemaattinen, mistä johtuen sen luotettavuus on vähäinen. State-of-the-art review on yleiskatsauksen alatyyppejä, joka keskittyy ajankohtaiseen tietoon ja asiakokonaisuuksiin. Tällainen katsaus tuo yleensä tiukan ajallisen poissulkukriteerin tiedonhakuun. [23 s. 12.]

### 5.1.2 Systemaattiset kirjallisuuskatsaukset

Integroiva kirjallisuuskatsaus on katsaustyypeistä laajin, ja sen tarkoituksena on tuottaa uutta tietoa, auttaa kriittisessä arvioinnissa ja tarkastelussa sekä synteettisessä luomisessa. Integroiva katsaus voi olla osa systemaattista katsausta, ja siihen voidaan sisällyttää kuvailevan katsauksen osia ja piirteitä. Tuloksena tällaisella katsauksella on laaja ja syvä synteesi aiheeseen. [23, s. 13.]

Systemaattinen katsaus on kirjallisuuskatsauksista tunnetuin ja käytetyin, sillä se on järjestelmällinen, syvälinen ja tuottaa tarkkaa tietoa. Tiedonhaketapa on laaja ja tarkka, ja katsaus sisältää usein meta-analyysin tai sellaisen pohjan. Systemoitu katsaus on systemaattisen katsauksen alatyyppejä, mutta suppeampi kaikin tavoin. Yleensä vain yksi tutkija suorittaa haun yhteen tietokantaan, jolloin myös tulokset ovat epäluotettavampia. [23, s. 14]. Osa muistakin katsaustyypeistä on omaksunut systemaattisen katsauksen luoman tiedonhaketavan, mutta ne käsittelevät aineistoa eri tavoin.

Järjestelmällinen haku ja katsaus (systematic search and review) jäljittelee myös systemaattista katsausta, mutta ei paneudu analysointiin ja laadunarviointiin yhtä voimakkaasti. Tutkimuskysymys on usein laajempi, mutta myös tulokset ovat vaihtelevia, sillä mukaanotto- ja poissulkukriteerit eivät ole yhtä tarkkoja. [23, s. 15.]

Sateenvarjokatsaus (umbrella review) on systemaattisten katsausten katsaus. Tällaiset katsaukset ovat harvinaisempia, sillä ne vaativat useampia systemaattisia katsauksia samasta tai samantapaisista aiheista. Tällainen katsaus tehdään yleensä aiheeseen, joka perustuu näyttöön, sillä kahden tai useamman tutkimuksen otanta on laajempi kuin yhden, ja näin tulosten vertailu on helpompaa ja laajempaa. Ongelmana sateenvarjokatsauksessa on, että tekijän tulee pystyä luottamaan kaikkeen lähdemateriaaliin. [23, s.15.]

Lisäksi määrällisiin tutkimuksiin perustuvissa systemaattisissa katsauksissa voidaan käyttää apuna meta-analyysia. Meta-analyysin tarkoituksena on luoda tilastollinen synteesi alkuperäistutkimuksista ja parantaa katsauksen luotettavuutta. [23 s. 17.]

## 5.2 Scoping-katsaus

Kirjallisuuskatsauksen tyyppiä valittiin kuvaileva scoping-katsaus, joka perustuu Arksey & O'Malley'n [24] tekemään työhön ja siihen, mitä Levac ym. [25] myö-

hemmin laajensi. Scoping-katsauksen tarkoituksena on luoda nopeasti ymmärrys aiheeseen, tuottaa tietoa strategiseen päätöksentekoon ja synteessin tekemiseen jatkokäyttöön. Scoping-katsauksella voidaan myös tunnistaa tutkimustiedossa olevia puutteita. Heikkoutena menetelmällä on, että se ei ota kantaa pohjatutkimuksien laatuun, ja sitä on vaikea toteuttaa, jos aiheesta ei ole riittävästi aikaisempia tutkimuksia. Yksin kirjallisuuskatsausta tehtäessä tulee myös varmistaa tutkijan motivaatio aiheeseen, jotta katsauksen eri vaiheissa ei aleta oikomaan. [23, s. 10–11; 26, s. 31.]

Scoping-katsausta voidaan kuvata nopean ja kartoittavan [24; 27] katsauksen yhdistetyksi alatyypiksi, sillä tarkoituksena on nopeasti hahmottaa laaja alue, ja tutkimuskysymyksen tarkkuus on anteeksiantavampi kuin systemaattisessa katsauksessa [28]. Scoping-katsaus valittiin katsaustyyppiksi, koska systemaattinen katsaus vaatii vähintään kaksi alan tuntevaa tutkijaa [23, s. 27; 29], mikä ei tässä opinnäytetyössä ole mahdollista. Kahden tutkijan vaatimuksessa jouduttiin joustamaan toisenkin kerran, sillä scoping-katsauksen järjestelmällinen tiedonhaku ja aineiston valinta vaatisi myös vähintään kaksi tutkijaa [30]. Valinnassa kuitenkin pysyttiin, sillä järjestelmällinen tiedonhaku haluttiin pitää oppimismahdollisuutena, vaikka työn luotettavuus madaltui.

Taulukossa 1 on lueteltuna scoping-katsauksen eri vaiheet, joista viimeinen eli konsultointi, vaihe kuusi (6), on Arksey & O'Malley'n [24] pohjakehyksen mukaan valinnainen. Levac ym. [25] on kuitenkin eri mieltä, sillä heidän laajennetun kehyksen mukaan sen tulisi olla keskeinen osa scoping-katsausta. Konsultoinnin tarkoituksena on informoida osakkeenomistajia, yrityksen johtoa ja asiantuntijoita, sekä näyttää mahdollista alustavaa tietoa. Tämä helpottaa myös tiedonvaihtoa ja voi antaa tutkijoille uusia näkökulmia.

Taulukko 1. Scoping-katsauksen vaiheet 1–6 [24].

Vaihe	Sisältö
1.	Identifioidaan tutkimuskysymys käyttäen laajoja käsitteitä.
2.	Identifioidaan asiaan liittyvät tutkimukset.
3.	Tutkimuksien ja kirjallisuuden valinta: ensin otsikoiden, sitten tiivistelmien perusteella. Kun artikkelit on valittu tiivistelmien perusteella, luetaan ne kokonaan ja poistetaan aiheeseen liittymättömät.
4.	Tiedon erottelu artikkeleista valmiin suunnitelman ja taulukon avulla, jotta keskitytään oikeisiin asioihin.
5.	Yhteenveto ja tulosten raportointi.
6.	Konsultointi, haastattelu ja informointi.

Mielestäni Levac ym. [25] laajennetun kehyksen tärkeimmät asiat liittyvät vaiheisiin 3, 4 ja 6. Vaiheessa 3 Arksey & O'Malleyn kehys antaa vaatimattoman lineaarisen kuvan tiedonhaun ja tutkimuksien valinnasta. Levac ym. [25] mukaan vaihetta pitäisi lähestyä iteratiivisesti tehden uusia hakuja aina, kun uutta tietoa löytyy. He ehdottavat myös samaa iteratiivista lähestymistapaa tiedon erotteiluun vaiheessa 4, jossa erottelutapaa ja hakutaulukkoa päivitetään jatkuvasti. Tämä lisäisi tulosten tarkkuutta ja varmistaisi aiheessa pysymisen sekä tutkimuskysymyksen oikeellisuuden. Levac ym. [25] parannusehdotuksissa puhutaan scoping-katsauksen tekemisestä vähintään kahden tutkijan tiimeissä, mikä ei tämän opinnäytetyön puitteissa ole mahdollista.

### 5.3 Tiedonhaun teoria

Järjestelmällisessä tiedonhaussa käytetään ennalta määrättyä strategiaa, sillä virheet haussa voivat viedä työltä luotettavuuden ja vääristää johtopäätöksiä. Työn alkaessa suoritetaan alustavia testihakuja, jotta saadaan käsitys kirjallisuuden määrästä, ja voidaan muokata hakusanoja ennen varsinaista aloitusta. [23, s. 25.] Jokainen haku tallennetaan, jotta siihen voi palata myöhemmin ja helpotetaan toistettavuutta. Haut suoritetaan kolmessa tasossa, jotka ovat otsikko, abstrakti ja kokoteksti. Jokaiselta tasolta tehdään kattavat muistiinpanot

ja taulukko, jotta seuraaminen helpottuu. Taulukot mahdollistavat niin oman kuin lukijankin hakuprosessin toistamisen ja parantavat luotettavuutta [23, s. 27].

Töiden valintojen jälkeen voidaan tehdä lisäksi manuaalinen haku esimerkiksi töissä käytettyjen lähteiden pohjalta, joissa voi olla hakukoneiden ulkopuolelle jäänyttä laadukasta tietoa [23, s. 27]. Lähteet voivat sisältää sanoja, aiheita tai tietokantoja, joita tekijä ei ole löytänyt tai joiden olemassaolosta ei tiedetty.

Työn viimeisessä vaiheessa tiedot järjestetään ja niistä tehdään yhteenveto. Ristiriitaiset tulokset viittaavat jatkotutkimuksen tai systemaattisen kirjallisuuskatsauksen tarpeeseen. [23, s.31.]

Hakusanoiksi ideoidaan mahdollisimman paljon erilaisia sanoja ja lauseita, käännetään ne englanniksi ja mietitään synonyymejä sekä aiheeseen liittyviä sanoja ja asioita [23, s.37]. Apuna voidaan käyttää esimerkiksi sanakirjoja, Wikipediaa ja tietokantoja. Testihakujen perusteella löydettyjen tutkielmien abstrakteista voi löytyä töissä yleisesti käytettyjä sanoja, joita voidaan hyödyntää. Hakusanat taulukoidaan ja niitä kokeillaan eri muodoissa, sillä eri osissa maailmaa ja eri kirjoittajien kesken sanoja on voitu muotoilla eri tavoin. Pelkästään britti- ja amerikanenglannin välillä on eroja, jotka voivat tuottaa erilaisia tuloksia tietokannoista. [23, s.37–38.]

Yleisimmin käytetyt tietokannat tukevat Boolean operaattoreita, joita ovat AND, OR ja NOT. AND-operaattori yhdistää asiasanoja toisiinsa, eli hakutuloksessa tulee olla molemmat mainitut sanat. OR-operaattorilla voidaan etsiä toisiinsa liittyviä tuloksia ja synonyymejä. NOT-operaattori poistaa asiasanojen esiintymistä, ja sen kanssa tulee olla tarkkana, sillä englannin kielessä kielteisyys ei välttämättä näy sanassa itsessään ja voi poistaa hyviä tuloksia. Operaattoreiden tärkeysjärjestystä ja hakua muokataan käyttämällä sulkeita. Lisäksi voidaan käyttää katkaisumerkkiä, yleensä asteriski (\*), jolloin haetaan sanojen osilla ilman, että tarvitsee eritellä päätteitä. Suomenkielisillä sanoilla haettaessa tulee olla tarkka, sillä sananvartalot usein taipuvat jo ennen päätettä. [23, s. 39–41.]

Hakutermit PICO [23, s. 18; 31] tulee sanoista Patient, Intervention, Comparison, Outcome (potilas, interventio, vertailuinterventio, hoidon tulos). Se on terveydenhuollon alalla paljon käytetty pohja, sillä se auttaa muotoilemaan tutkimuskysymyksen kliniseen ongelmaan tarkentuen [32]. PICO-periaatetta ei käytetä tässä opinnäytetyössä, sillä kirjallisuuskatsauksessa ei paneuduta potilaisiin, tutkimusasetelmaan tai hoitojen tuloksiin: tarkoituksena on löytää teknologista tietoa.

#### 5.4 Haku- ja valintakriteerit

Tutkimuksien valintakriteereissä jouduttiin joustamaan poiketen systemaattisen kirjallisuushaun suosituksista [23, s. 26; 33]. Kielet rajattiin suomen ja englannin kieleen ymmärrettävyyden takia. Lisensoitua materiaalia haettiin ja käytettiin niistä tietokannoista, joihin Metropolia ammattikorkeakoululla oli voimassa oleva lisenssi. Aikarajaus tehtiin manuaalisen testihaun yhteydessä, kun huomattiin, että vuodesta 2010 eteenpäin aiheeseen liittyvien artikkelien määrä on lähtenyt nopeaan kasvuun. Mukaan otettiin vain käytetyistä tietokannoista löytyvä julkaistu materiaali, joiden otsikko ja tiivistelmä liittyivät aiheeseen. Sisäänotto- ja poissulkukriteerit on eritelty taulukossa 2.

Taulukko 2. Tietokantahakujen sisäänotto- ja poissulkukriteerit.

<b>Sisäänottokriteerit</b>	<b>Poissulkukriteerit</b>
Julkaisu vuosina 2010–2021	Julkaistu ennen vuotta 2010
Materiaalin kieli suomi tai englanti	Materiaalin kieli ei ole suomi tai englanti
tutkimusartikkelit, alan julkaisut, muut kirjallisuuskatsaukset	opinnäytetyöt, ei julkaistu materiaali
Ilmainen, vapaa ja lisensoitu materiaali, mihin Metropolia Ammattikorkeakoululla on lisenssi	Materiaali mihin Metropolia Ammattikorkeakoululla ei ole lisenssiä
Otsikko liittyy aiheeseen	Otsikko ei liity aiheeseen
Tiivistelmä liittyy aiheeseen	Tiivistelmä ei liity aiheeseen

### 5.5 Aineiston etsiminen

Ennen varsinaista hakua suoritettiin lyhyt manuaalinen testihaku ja tarkistettiin kahdesta löydetyistä (n=2) englanninkielisestä tutkimuksesta niissä käytettyjä sanoja ja muotoja, jotta hakusanastoa voitiin tarkentaa. Hakusanat on eritelty taulukossa 3 suomen- ja englanninkielisiin sarakkeisiin. Tarkemmat sanamuodot, katkaisut sekä yhdisteleminen on esitetty luvussa aineiston valinta.

Taulukko 3. Tietokantahauissa käytetyt suomen- ja englanninkieliset hakusanat.

<b>Suomenkieliset hakusanat</b>	<b>Englanninkieliset hakusanat</b>
tilastollinen, ennakointi, mallinnus, kapasiteetti, hoitajakso, ennustaminen, potilasvirta	predict, anticipation, forecasting, bed, patientflow, length of stay, occupation, predictive model, hospital, management, shortage, patient



Joitakin tietokantahakuja suoritettiin iteratiivisesti hakusanojen kehittyessä, kuten Levac ym. [25] suosittaa kehyksessään, mutta kaikkia hakuja ei aikarajoitusten vuoksi voitu uusaa. Tiivistelmien perusteella töitä otettiin mukaan kevyin perustein, yleensä muutaman kiinnostavan tai asiaan liittyvän sanan perusteella. Kaikki haut rajattiin vuosien 2010 ja 2021 välille PubMed:n testihaun perusteella. Käytetyt tietokannat ja niiden kuvaukset ovat taulukossa 4. Haut suoritettiin 15.7 – 20.8.2021 välisenä aikana.

Taulukko 4. Käytetyt tietokannat ja niiden kuvaukset.

Tietokanta	Kuvaus
PubMed	Yhdysvaltalainen hakukone ja viitetietokanta
CINAHL	Kansainvälinen hoitotieteen viitetietokanta. Metropolia Ammattikorkeakoulun lisenssi.
IEEEExplore	Kansainvälinen sähkö- ja tietotekniikan kokotekstitietokanta. Metropolia Ammattikorkeakoulun lisenssi.
MEDIC	Suomalainen terveystieteellinen viitetietokanta. Metropolia Ammattikorkeakoulun lisenssi.
JBI Ovid	Australialainen näyttöön perustuvan hoitotyön tietokanta. Metropolia Ammattikorkeakoulun lisenssi.
ScienceDirect	Kansainvälinen monitieteellinen tietokanta. Metropolia Ammattikorkeakoulun lisenssi.

## 5.6 Valitut aineistot

Ensimmäinen haku tehtiin PubMed-tietokantaan ja julkaisuajankohdaksi määritettiin 2010–2021. Hakusanoina käytettiin ("predicting" OR "anticipation") AND ("bed occupation" OR "patientflow" OR "patient flow") ja löydettiin 18 osumaa käyttäen asiasanahakua ja valittuna "maksuton kokoteksti". Mukaan otettiin 8 otsikkovalinnassa ja niistä 6 abstraktin perusteella. Koska heti ensimmäisen testihaun aikana huomattiin, että hoitojakson pituus mainittiin useasti, otettiin se mukaan hakuun. Toisessa haussa hakusanoina käytettiin ("predictive model" OR "forecasting") AND "length of stay", ja näillä saatiin 181 osumaa, joista 21

otettiin mukaan otsikkovalinnassa ja niistä 13 valikoitui lopuksi tiivistelmävalinnassa. Kolmas haku vastasi lähes täysin edellistä, mutta hoitojakson pituus (Length of stay = LoS) jätettiin pois vinouman (bias) vähentämiseksi, sillä hoitojakson pituutta voidaan ennustaa itsessään, tai sitä voidaan käyttää parametrinä potilasmäärien ennustamisessa. Hakusanoina oli ("predictive model" OR "forecasting") AND ("bed" OR "occupation"), ja se tuotti 151 tulosta, joista 21 otettiin mukaan otsikkovalinnassa ja niistä 9 jäi abstraktien perusteella. Viimeisessä haussa lähestymistapaa muutettiin otannan laajentamiseksi, ja käytettiin uusia hakusanoja "hospital bed shortage" OR "hospital bed management". Tämä tuotti 22 osumaa, joista 13 otettiin mukaan otsikon perusteella, ja niistä 5 abstraktien perusteella. Hakusanat on eritelty taulukossa 5.

Taulukko 5. PubMed-haku.

ID	Hakusanat	Osumat n=	Otsikkovalinta n=	Abstrakti valinta n=
1	("predicting" OR "anticipation") AND ("bed occupation" OR "patientflow" OR "patient flow")	18	8	6
2	("predictive model" OR "forecasting") AND "length of stay"	181	21	13
3	("predictive model" OR "forecasting") AND ("bed" OR "occupation")	151	21	9
4	"hospital bed shortage" OR "hospital bed management"	22	13	5
	Yhteensä	372	63	33

Toinen haku suoritettiin CINAHL-tietokantaan ja tulokset on rajattu ilmaisiin kokoteksteihin. Tämän lisäksi CINAHL-haku synonymisoi joitakin sanoja, jotka vaikuttivat osumien määrään. Hakusanat on eritelty taulukossa 6. Ensimmäisessä

haussa hakusanoina oli ("predicting" OR "anticipation") AND ("bed occupa\*" OR "patientflow" OR "patient flow"). PubMed-hakuun verrattuna lisättiin tässä yksi katkaisumerkki laajentamaan hakua. Osumia oli 37, joista otsikon perusteella valittiin 5 ja niistä abstraktien perusteella 3 lopulliseen valintaan. Toisen haun hakusanat olivat ("predictive model" OR "forecasting") AND "length of stay", jotka tuottivat 222 osumaa, joista 12 otettiin mukaan otsikon perusteella, ja niistä 9 tiivistelmien perusteella. Kolmannessa haussa hakusanoina oli ("predictive model" OR "forecasting") AND ("bed" OR "occupation"), jotka tuottivat 122 osumaa, ja joista vain 4 valittiin otsikon perusteella, ja niistä 3 abstraktien perusteella. Viimeisen haun hakusanoina oli "hospital bed shortage" OR "hospital bed management", joilla löytyi 10 osumaa. Otsikoiden perusteella mukaan otettiin 2, joista vain toinen valittiin abstraktin perusteella.

Taulukko 6. CINAHL-haku.

ID	Hakusanat	Osumat n=	Otsikko-valinta n=	Abstrakti valinta n=
1	("predicting" OR "anticipation") AND ("bed occupa*" OR "patientflow" OR "patient flow")	37	5	3
2	("predictive model" OR "forecasting") AND "length of stay"	222	12	9
3	("predictive model" OR "forecasting") AND ("bed" OR "occupation")	122 <sup>1</sup>	4	3
4	"hospital bed shortage" OR "hospital bed management"	10	2	1
	Yhteensä	269	23	16
	(1 CINAHL-haku synonymisoi sanat "occupation" ja "employment", mistä johtuu otsikkovalintojen pieni lukumäärä)			

Kolmas haku tehtiin IEEEXplore-tietokantaan, joka eroaa kahdesta aiemmasta tietokannasta keskittyen tietoteknisiin ja teknologisiin artikkeleihin ja tutkimuksiin

lääketieteellisten ja terveydenhuollollisten sijaan. IEEEExplore valittiin mukaan mahdollistamaan erilainen lähestymistapa terveydenhuollon ongelmaan. Hakusanat on eritelty taulukossa 7. Ensimmäisessä haussa hakusanoina oli ("predicting" OR "anticipation") AND ("bed occupa\*" OR "patientflow" OR "patient flow"), joka tuotti 10 osumaa, joista 5 otettiin mukaan otsikkovalinnassa, ja niistä 4 abstraktien perusteella. Toinen haku tehtiin hakusanoilla ("predictive model" OR "forecasting") AND "length of stay", joka tuotti 14 osumaa, joista 4 valittiin otsikon perusteella, ja kaikki kyseiset myös abstraktien perusteella. Kolmas haku tehtiin hakusanoilla ("predictive model" OR "forecasting") AND ("bed" OR "occupa\*"), ja se tuotti 254 osumaa, joista vain 3 otettiin mukaan otsikon perusteella, ja niistä 2 abstraktin perusteella. Neljäs ja viides haku jouduttiin aikaisemmista poiketen erottamaan OR-operaattorin toiminnan johdosta. Neljäs haku tehtiin hakusanalla "hospital bed management", ja se tuotti 131 osumaa, joista 11 valittiin otsikon perusteella, ja niistä 8 abstraktin perusteella. Viides haku suoritettiin hakusanalla "hospital bed shortage", joka tuotti 15 osumaa. Otsikon perusteella mukaan otettiin vain 1, joka jäi pois abstraktin lukemisen jälkeen.

Taulukko 7. IEEEXplore-haku.

ID	Hakusanat	Osumat n=	Otsikko-valinta n=	Abstrakti valinta n=
1	("predicting" OR "anticipation") AND ("bed occupa*" OR "patientflow" OR "patient flow")	10	5	4
2	("predictive model" OR "forecasting") AND "length of stay"	14	4	4
3	("predictive model" OR "forecasting") AND ("bed" OR "occupa*")	254	3	2
4	"hospital bed management"	131 <sup>1</sup>	11	8
5	"hospital bed shortage"	15	1	0
	Yhteensä	293	24	18
	(1 OR-operaattorin erilainen käyttäytyminen, aikaisempien hakujen ID 4 hajautettiin kahdeksi.)			

Neljäs haku tehtiin suomenkieliseen MEDIC-tietokantaan, jossa synonyymien käyttö voitiin valita erikseen ja se pidettiin valittuna koko haun ajan. Hakusanoina oli (tilastollinen OR ennakointi OR mallinnus) OR kapasiteetti OR hoitajakso OR ennustaminen OR potilasvirta (hakusanat eroteltuna taulukossa 8), joilla osumia tuli 86. Otsikkovalinnassa mukaan otettiin 12, joista abstraktien perusteella valittiin 2.

Taulukko 8. MEDIC-haku.

ID	Hakusanat	Osumat n=	Otsikko-valinta n=	Abstrakti valinta n=
1	tilastollinen OR ennakointi OR mallinnus	42	6	1
2	kapasiteetti	5	2	0
3	hoitojakso	28	2	0
4	ennustaminen	10	1	1
5	potilasvirta	1	1	0
	Yhteensä	86	12	2
	(*synonyymit käytössä koko haun ajan)			

JBI Ovid on kliininen tietokanta, ja vaikka osumia oli, eivät ne liittyneet tutkittavaan aiheeseen, eikä tietokannasta otettu mukaan yhtäkään artikkelia. Hakusanoilla ("predicting" OR "anticipation") AND ("bed" OR "occupancy" OR "occupation" OR "patient flow") saatiin tuloksia 16. Hakusanoilla ("predictive model" OR "forecasting") AND "length of stay" saatiin tuloksia 63. Hakusanoilla ("predictive model" OR "forecasting") AND ("bed" OR "occupa\*") saatiin tuloksia 16. Kaksi viimeistä hakua erotettiin monikenttähauksi: (hospital and bed and management) tuotti 39 osumaa, ja (hospital and bed and shortage) tuotti 4 osumaa. JBI Ovid -haku on kuvattu taulukossa 9.

Taulukko 9. JBI Ovid -haku.

ID	Hakusanat	Osumat n=	Otsikko-valinta n=	Abstrakti valinta n=
1	("predicting" OR "anticipation") AND ("bed" OR "occupancy" OR "occupation" OR "patient flow")	16	0	0
2	("predictive model" OR "forecasting") AND "length of stay"	63	0	0
3	("predictive model" OR "forecasting") AND ("bed" OR "occupa*")	16	0	0
4	(hospital and bed and management) <sup>1</sup>	39	0	0
5	(hospital and bed and shortage) <sup>1</sup>	4	0	0
	Yhteensä	138	0	0
	( <sup>1</sup> Monikenttähaku, hakusanat eroavat muista.)			

Viimeisenä tehtiin haku ScienceDirect-tietokantaan, joka erosi muista tietokannoista ollessaan huomattavasti laaja-alaisempi, eikä keskittynyt pelkästään lääketieteellisiin tai terveydenhuollollisiin artikkeleihin. ScienceDirect-haussa myös hyödynnettiin paljon lisärajoituksia aiheeseen liittyen, ja haut tehtiin otsikkohakuna. Hakusanat ja lisärajoitukset on eritelty taulukossa 10. Ensimmäisen haun hakusanat olivat (predict OR anticipation OR forecasting) AND (bed OR "patient flow" OR "length of stay" OR occup), joita tietokanta osasi täydentää itsenäisesti. Tämä tuotti 318 osumaa, joista 15 otettiin mukaan otsikon perusteella ja joista 10 valittiin abstraktin perusteella. Toisen haun hakusanat olivat "predictive model" AND "patient flow" OR "hospital bed". Näillä löydettiin 53 osumaa, joista 7 otettiin mukaan otsikon perusteella, ja niistä 5 abstraktin perusteella. Kolmas haku muodostui hakusanoista hospital AND bed AND management, joka tuotti 2 osumaa ja joista yhtäkään ei valittu. Neljäs haku oli lähes identtinen kolmannen kanssa, hakusanat olivat hospital AND bed AND shortage, ja ne tuottivat

myös 2 osumaa, joista yhtäkään ei valittu. Viimeisessä haussa hieman iteroitiin ja Levac ym. [25] suosituksen mukaan tehtiin yksi uusi haku hakusanoilla "predictive model" AND patient, joka tuotti 171 osumaa. Niistä vain 2 otettiin mukaan otsikon perusteella ja vain toinen valittiin tiivistelmän perusteella.

Taulukko 10. ScienceDirect-haku.

ID	Hakusana	Osumat n=	Otsikko-valinta n=	Abstrakti valinta n=
1	(predict OR anticipation OR forecasting) AND (bed OR "patient flow" OR "length of stay" OR occup) <sup>1</sup>	318	15	10
2	"predictive model" AND "patient flow" OR "hospital bed" <sup>2</sup>	53	7	5
3	hospital AND bed AND management <sup>3</sup>	2	0	0
4	hospital AND bed AND shortage <sup>3</sup>	2	0	0
5	"predictive model" AND patient <sup>4</sup>	171	2	1
	Yhteensä	546	24	16
	(1 Lisärajaus: Medicine and Dentistry, Engineering, Nursing and Health Professions, otsikkohaku) (2 Lisärajaus: Medicine and Dentistry, Nursing and Health Professions, Social Sciences, Engineering, Computer Science, Decision Sciences, Mathematics, otsikkohaku) (3 Lisärajaus: otsikkohaku) (4 Lisärajaus: Medicine and Dentistry, Computer Science, Nursing and Health Professions, Social Sciences, Engineering, otsikkohaku)			

Tietokannoissa oli jonkin verran päällekkäisyyksiä, joten duplikaatit tarkistettiin ja poistettiin käyttäen RefWorks-viitteidenhallintaohjelmaa [23, s. 53] Metropolia

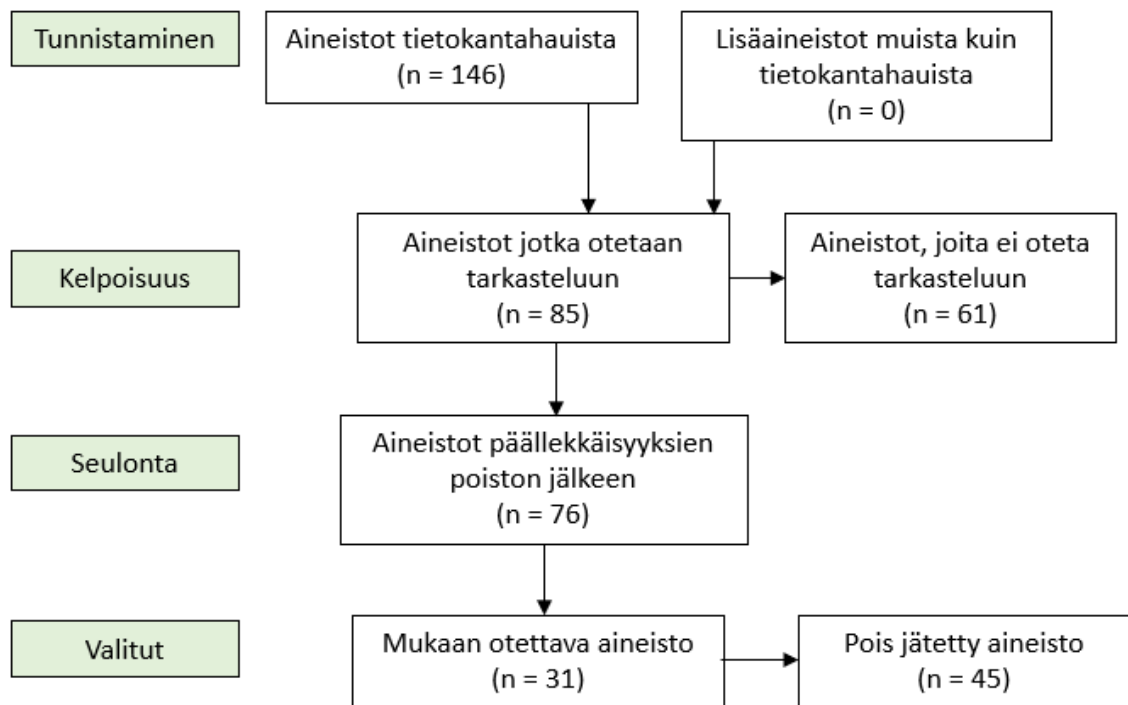


ammattikorkeakoulun lisenssillä. Kuvassa 4 on esitettyä hakuprosessi vuokaviona, joka mukaillee Stoltin ja Moherin [23, s. 63] mallia, mutta eri järjestyksessä. Tässä työssä duplikaatit poistettiin vasta tiivistelmävalinnan yhteydessä.

Otsikkovalinnassa mukaan otettiin yhteensä 146 artikkelia, eikä hyödynnettävää lisäaineistoa löytynyt tietokantahaun aikana. Valituista artikkeleista luettiin abstraktit, joiden perusteella kokotekstivalintaan jäi 85 artikkelia ja pois jätettiin 61 artikkelia. Koska päällekkäisyydet poistettiin vasta tässä vaiheessa, jäi analysoitavaksi 76 artikkelia. Kaikki mukana olleet artikkelit ovat listattuna liitteessä 1. Kokotekstianalyysi suoritettiin 4.9 - 16.9.2021 välisenä aikana, ja lopulta 31 artikkelia koodattiin [34] ja dokumentoitiin. Muut kokotekstianalyyssissä mukana olleet artikkelit (n = 45) jätettiin ulkopuolelle erinäisistä syistä.

- Ennakoitiin väärää asiaa (esim. vuoteiden kokonaismäärää tai COVID-19 leviämisenopeutta).
- Ennakoitiin väärää aikaväliä (esim. kuinka paljon vuoteita tarvittaisiin yhteensä 5 vuoden päästä).
- Ennakoivan mallin käyttöalue tai suuntaus oli aiheeseen liittymätön (esim. ennakoitiin potilasmäärää maantieteellisessä alueella).
- Tutkimuksessa ei kerrottu käytettyjä metodeja tai parametreja tai tiedot olivat puutteelliset.
- Tutkimuksen tulokset olivat epäselvät.
- Tutkimus keskittyi taloudelliseen tai matemaattiseen malliin tai ei-ennakoivaan metodiin (esim. johtamistapa).

Edellä mainituissa tehtiin muutama poikkeus, jos tutkimus sisälsi muuten hyödynnettävää dataa tai mahdollisesti käytettäviä parametreja. Kaksi tutkimusta otettiin mukaan esimerkkinä erilaisista malleista (COVID-19 leviäminen n=1, matemaattinen malli n=1).



Kuva 4. Vuokaavio aineiston valinnan etenemisestä Stolt ym. (2016) ja Moher (2009) mukailleen [23, s. 63].

## 5.7 Aineiston arviointi

Tässä opinnäytetyössä ei otettu kantaa aineiston laatuun, luotettavuuteen tai sen arviointiin muilta osin kuin hakemalla aineisto tässä opinnäytetyössä kuvatuista lähteistä. Toisin kuin systemaattisessa kirjallisuuskatsauksessa, laadun tai luotettavuuden arviointi ei ole scoping-katsauksen tarkoitus. Systemaattisen kirjallisuuskatsauksen tiukkojen laadunarviointikriteerien vuoksi paljon tutkimuksia jää analysoinnin ulkopuolelle, tai niille annettu painoarvo on vähäinen. Scoping-katsauksen tavoitteena on tutkimusten laaja otanta, analysointi ja tulosten kuvaileva esittäminen. Tämä on kuitenkin yksi katsauksen heikkouksista, sillä suuren materiaalmäärän esittäminen voi olla vaikeaa. [24.]

## 5.8 Aineiston analysointi

Aineiston analysointi aloitettiin tiivistelmävalinnan ja päällekkäisyyksien poiston jälkeen lukemalla kaikki kokotekstit (n = 76). Kokotekstien analysointi suoritettiin

laadullisen sisällönanalyysin keinoin [34], pääasiassa koodausmenetelmällä. Koodaus tarkoittaa tekstin erottelua pienempiin osiin yksinkertaistaen sitä tutkijan määrittelemiin luokkiin, tässä opinnäytetyössä taulukon sarakkeisiin. Koodaaminen voi olla joko teoria- tai aineistoperusteista; teoriaperusteisessa koodaamisessa luokat tai kategoriat ovat ennalta olemassa, kun taas aineistoperusteisessa kategorioita luodaan aineiston perusteella. [35.] Tässä opinnäytetyössä käytettiin molempia koodausperiaatteita. Analysointi aloitettiin teoriaperusteisesti, ja luokat olivat valmiina jo ensimmäisen tutkimuksen kohdalla. Levac ym. [25] scoping-katsauksen parannuskehityksen johdosta analysointi suoritettiin iteratiivisesti, ja tutkimukset luettiin kaksi kertaa. Aineistoperusteinen koodaus otettiin mukaan toisella lukukerralla, ja luokkia luotiin yksi lisää.

Koodaus tehtiin Excel-taulukoon ja luokkia syntyi lopulta 11:

- otsikko
- tekijät
- julkaisuvuosi
- tutkimustyyppi
- otannan suuruus (tutkimuksesta riippuen henkilöä tai vastaanottoa)
- otannan kesto (kuukausina)
- mallin tyyppi
- algoritmit
- parametrit
- ennakkoinnin tarkkuus ja pituus
- suosittelevatko tutkijat käyttöä.

Vaikein luokiteltava asia oli tutkimustyyppi, sillä se oli kerrottu vain yhdeksässä tutkimuksessa. Toiseksi vaikein luokiteltava asia oli mallin tyyppi, sillä sitä ei kerrottu kaikissa tutkimuksissa. Joissakin tutkimuksissa mallin tyyppin pystyi päättämään käytetyistä algoritmeista, esimerkiksi satunnaismetsämalli on pelkästään koneoppimismenetelmä. Tutkimuksista 13 käsitteli koneoppivia malleja, 4 tilastollisia malleja, ja yhdistettyjä, muita tai ei määriteltyjä malleja oli 14. Taulukko kaikista tuloksista ja luokista on esiteltyinä liitteessä 2.

Otantojen suuruudet vaihtelivat, eivätkä kaikki tutkimukset sisältäneet tietoa käytetystä potilasdatan määrästä. Pienimmässä otannassa oli 104 potilaan tiedot ja laajimmassa 715 143 potilaan tiedot. Otantojen suuri vaihteluväli selittyy erilaisilla tutkimusympäristöillä ja tutkimuksen laajuudella. Pienimmässä otannassa oli mukana yksi päivystysosasto, suurimmassa Yhdysvaltojen koko NSQIP-tietokanta (National Surgical Quality Improvement Program) [36]. Otantojen kestot vaihtelivat kahden kuukauden ja yhdeksän vuoden välillä.

Algoritmeista yleisimpiä olivat erilaiset satunnaismetsä- (random forest, RF) ja regressiomallit, sekä neuroverkot (artificial neural network, ANN). Vaikka regressiomallit ovat yleisesti tilastotieteellisiä, käyttävät monet koneoppimismenetelmät niitä hyväkseen [37].

Tutkimuksissa käytetyimmät parametrit olivat normaalit potilastiedot kuten ikä, sukupuoli ja tulon syy tai alkuperäinen diagnoosi, mitkä esiintyivät yli 2/3:ssa tutkimuksista. Normaaliin potilastietojen lisäksi tutkimuksissa käytettiin paljon erilaisia koodeja tai luokitusjärjestelmiä osastosta, sairaalasta, maantieteellisestä alueesta tai tutkittavasta aiheesta riippuen. Esimerkiksi lonkkamurtumaleikkauksen jälkeisen hoitojakson pituutta ennakoivissa tutkimuksissa (n = 3) oli yhtenä parametrina ASA-luokitus (American Society of Anesthesiologists), eli fyysisen tilan luokitusjärjestelmä (physical status classification system) [38].

Teho- ja päivystysosastoilla tehdyissä tutkimuksissa käytettiin useita erilaisia luokituksia, pisteytyksiä ja koodeja, esimerkiksi Glasgow Coma Scale (GCS), Sequential Organ Failure Assessment (SOFA) score sekä Early warning score (EWS). Tulon syy tai alkuperäinen diagnoosi oli usein merkitty ICD-9-koodilla. Muita käytettyjä koodeja sekä kaikki tutkimuksissa käytetyt parametrit ovat liitteen 2 taulukossa rivillä ”parametrit”. Lisäksi Kabir S. & Farrokhvar L. [39] tuottamassa epälineaarisen parametrivalinnan mallissa mukana olleet parametrit ja tulokset ovat liitteissä 3 ja 4.

Stecker M. ym. [40] tekemässä tutkimuksessa keskityttiin etsimään pelkästään parametreja, joilla on suuri painoarvo ennakoivissa malleissa neurotieteiden

osastoilla. Kyseissä tutkimuksessa ei arvioitu mallin tarkkuutta, vaan etsittiin mahdollista korrelaatiota eri parametrien ja hoitajaksojen pituuksien välillä. Tutkimuksen mukaan suurin painoarvo oli mRS-pisteytyksellä (Modified Rankin Score), matalalla systolisella verenpaineella, intubaatiolla sekä mahdollisella laskimotulpalla, virtsatieinfektiolla tai potilaan joutumisella teho-osastolle.

Tutkimuksissa ilmenneiden tarkkuuksien keskinäinen vertailu on vaikeaa, sillä tutkijoista, mallin tyypistä, ympäristöstä, käytetyistä parametreista ja algoritmeista riippuen ennakkoinnin tarkkuutta ilmoitettiin monella eri tavalla. Tarkastellessa liitteen 2 taulukoiden rivejä ”ennakkoinnin tarkkuus ja pituus” huomataan, että tarkkuuksia on ilmoitettu esimerkiksi muodoissa accuracy, average precision (AP), receiver operating characteristic (ROC), mean absolute percentage error (MAPE), mean squared error (MSE) sekä area under receiver operating characteristic (AUROC). Neljässä tutkimuksessa oli annettu lisäksi vaihteluväli, jos esitystapa sen mahdollisti.

Tarkkuuksien suora vertailu kierrettiin iteroimalla kokotekstianalyysi [25] ja tutkimuksista etsittiin tietoa siitä, että suosittelivatko tutkijat kehittämänsä mallin käyttämistä. Tutkimuskohtaiset tulokset ovat liitteen 2 taulukoiden riveillä ”suosittelevatko tutkijat käyttöä”. Tulokset jaettiin kolmeen vaihtoehtoon: ”Kyllä”, ”Ei” sekä ”Vaatii lisätutkimusta”. Mukana olleista 31 tutkimuksesta, 16:n tutkimuksen tekijät suosittelivat mallin käyttöä, ja 13:n tutkimuksen tekijöiden mielestä mallin käyttöönottoaminen vaatii lisätutkimusta. Vain yhden mallin kohdalla tutkijat eivät suosittaneet käyttöä, ja yhden tutkimuksen tarkoituksena ei ollut ennakoivan mallin arviointi tai luominen.

## **6 Tulokset**

### **6.1 Aineistosta saadut tiedot**

Yksittäisiä tutkimuksia tarkastellessa voidaan tuloksia hyödyntää tehokkaammin kuin keskenään vertaillen, kuten aineiston analysointi -luvussa mainittiin. Abd-

Elrazek ym. [41] tekemässä tutkimuksessa käytettiin 31 eri parametria ja parhaimmat tulokset saavutettiin kahdella eri koneoppimismenetelmällä. Sumealla logiikalla (fuzzy logic, FL) tarkkuudeksi saatiin noin 92 % ja päätöksentekopuu-algoritmillä (classification tree, CT) tarkkuus oli noin 90 %.

Barnes ym. [42] tekemässä tutkimuksessa ennakoitiin saman päivän kotiutumista koneoppimismenetelmillä. Regressiivisellä satunnaismetsämallilla (regression random forest, RRF) oli 82 %:n tarkkuus, kun ennakoitiin ennen klo 14:00 kotiutumista. Ennakoidessa päivän loppuun mennessä kotiutumista laski tarkkuus 63 %:iin. Tutkijat kuitenkin kertoivat, että joinakin päivinä malli pystyi ennakoimaan jopa järjestyksen, jossa potilaat kotiutuivat osastolta.

Vuonna 2021 julkaistussa hoitojakson pituutta ennakoivassa tutkimuksessa Harerimana ym. [43] tuottivat mallin, jolla pyrittiin vähentämään yksittäisten parametrien määrää. He käyttivät Hierarchical Attention Network (HAN) -koneoppimismenetelmää yhdessä luonnollisen kielen käsittelyjärjestelmän kanssa (Natural language processing, NLP), jotta he pystyivät hyödyntämään vapaita tekstikenttiä. Toisena pääasiallisena parametrina oli ensimmäisen 24 tunnin aikana annettu ICD-9-koodi. Mallin tarkkuus oli korkea, yli 90 %, ennakooidessa lyhyttä alle 10 päivän hoitojaksoa. Tarkkuus pysyi yli 90 %:ssa myös keskipitkän hoitojakson (10–30 päivää) ennakoinnissa, ja laski vain hieman alle 90 %:n pitkässä ennakoinnissa (yli 30 päivää). Tutkimuksen tarkoituksena oli saavuttaa korkea tarkkuus ennakoinnissa niillä tiedoilla, jotka potilaasta saatiin ensimmäisen 24 tunnin aikana.

Cereda ym. [44] tekemässä ikäihmisten ravitsemukseen ja sen tuomiin riskeihin liittyvässä tutkimuksessa esiteltiin korrelaatiota sairauden aiheuttaman aliravitsemuksen ja korkeamman kuolleisuuden sekä pidentyneen hoitojakson pituuden välillä. Tutkijat esittelivät GNRI:n (Geriatric Nutritional Risk Index) käyttämistä yhtenä ennakoivan mallin (tässä logistinen regressio) parametrinä. Indeksiksi yhdistää kaksi ravitsemuksellista indikaattoria [44]: seerumin albumiinin [45] ja henkilön nykyisen massan verrattuna ideaaliin massa [46]. GNRI:n laskukaava on esiteltyä alla kaavassa 1.

$$GNRI = \left( 1.487 \times \text{serum albumin}, \left[ \frac{g}{L} \right] \right) + \left( 41.7 \times \left( \frac{\text{present weight} [kg]}{\text{usual weight}} \right) \right) \quad (1)$$

*serum albumin: seerumin albumiini*

*present weight: nykyinen massa*

*usual (ideal) weight: ideaali massa*

*miehillä = 50 + (0.91 × (pituus [cm] − 152.4))*

*naisilla = 45.5 + (0.91 × (pituus [cm] − 152.4))*

Tutkimuksessa GNRI:n tuottamat luvut jaoteltiin kolmeen luokkaan, jossa alle 92 viittasi suureen riskiin, 92–98 pieneen riskiin, ja yli 98 ei riskiä. Tutkimuksen mukaan korkeampi riski pystyttiin yhdistämään pidempään hoitojaksoon, laihtumiseen sekä todennäköisyyteen, jossa potilaalla on useampia sairauksia. Tutkijat suosittelevat GNRI:n käyttämistä parametrina ennakkoinnissa myös akuutissa ympäristössä. [44.]

Koska Espoon sairaala on ikäihmisten hoitoon ja kuntoutukseen keskittyvä yksikkö, tulisi tulosten tarkastelu (liitteet 2, 3 ja 4) tutkimuskysymyksen perusteella aloittaa tutkimuksista, jotka keskittyvät esimerkiksi lonkkamurtumapotilaiden hoitojaksojen pituuksien ennakointiin. Lisäksi edellä mainittu Cereda ym. tekemä tutkimus GNRI:stä [44] sekä moniosastotutkimukset, esimerkiksi Kabir & Farrokhvar [39], sisältävät mahdollisesti hyödynnettäviä parametreja.

## 6.2 Konsultointi

Opinnäytetyön loppuvaiheen tiukan aikataulun vuoksi konsultointi suoritettiin kahtena lyhyenä keskusteluna Espoon sairaalan laatupäällikön Mari Järvisalon kanssa. Keskusteluissa käytiin läpi opinnäytetyön ja sen tuloksien sisältöä. Opinnäytetyön julkaisun yhteydessä Espoon sairaalalle lähetetään työssä tuotetut alkuperäistiedostot.

## 7 Yhteenveto

### 7.1 Tulosten tarkastelu ja pohdinta

Opinnäytetyössä tutkittiin terveydenhuollossa kokeiltuja ja käytössä olevia ennakoivia malleja. Työn tavoitteena oli luoda teoriapohja erilaisista malleista sekä niissä käytetyistä parametreista käyttäen scoping katsaus -menetelmää. Tutkimuksia analysoitaessa huomattiin, että ennakoiva malli pystytään toteuttamaan monella tavalla, ja mukana oli niin koneoppivia, tilastollisia kuin matemaattisia-kin malleja. Matemaattisten mallien analysointi ei kuitenkaan ollut mahdollista tämän opinnäytetyön aikana yhtä tutkimusta lukuun ottamatta. Scoping-katsauksessa mukana olleiden tutkimusten pääasiallisena tavoitteena oli ennustaa potilasmääriä, kotiuttamisaikoja ja hoitajaksojen pituuksia.

Työn tavoitteisiin päästiin ja tuloksena saatiin tutkimuksista taulukoitua erilaisia ennakoivia malleja ja niissä käytettyjä algoritmeja, parametreja sekä tarkkuuksia. Jos opinnäytetyö olisi toteutettu tiimityönä korkeammalla asteella, olisi tutkimuksista saatu irti todennäköisesti vielä enemmän, esimerkiksi ohjelmoinnin ja tilastomatematiikan osalta. Lisäksi olisi ollut mahdollisuus perehtyä matemaattisiin malleihin, laajentaa rajausta sekä hakea tietoa useammasta tietokannasta.

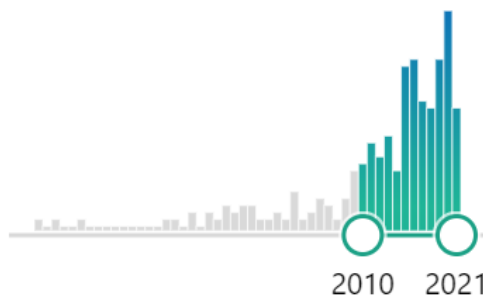
Analysoinnin aikana huomattiin, että yhtä täydellistä ennakoivaa mallia ei ole. Malli tulee aina optimoida jokaiselle osastolle erikseen, ja ensimmäisiin tuloksiin tulee suhtautua varauksella [14; 47]. Väitöstä vahvistaa myös Kabir & Farrokhtarin [39] tutkimus. Tarkasteltaessa liitettä 4 huomataan, että osastojenväliset tarkkuudet vaihtelevat, eivätkä tarkkuudet ole myöskään suhteessa lineaarisia verrattuna käytettyjen parametrien määrään, eli suurempi määrä ei takaa tarkempia tuloksia.

Covid-19-pandemia on vauhdittanut ennakoivien mallien tutkimusta ja kehitystä [48]. Tämä voidaan todeta kuvasta 5, joka selkeästi osoittaa vuonna 2020 PubMed-tietokannassa julkaistujen artikkelien määrän ylittävän aikaisempien vuosien julkaisumäärät. Sotilaslääketieteen aikakauslehdessä [49] oli myös artikkeli



puolustusvoimien käyttämästä SIR-mallinnuksesta (Susceptible, Infected, Recovered), jolla arvioidaan COVID-19-pandemian leviämistä. SIR-malleja ei kuitenkaan huomioitu tämän opinnäytetyön puitteissa muuta kuin osana yhtä tutkimusta, sillä ne olivat rajauksen ulkopuolella.

#### RESULTS BY YEAR



Kuva 5. PubMed-haun aikajana hakusanoilla: ("predictive model" OR "forecasting") AND "length of stay".

Tämä opinnäytetyö on erittäin ajankohtainen, sillä COVID-19 on aiheuttanut ruuhkia ja vuodepulaa ympäri maailman. Mediassa näkyvimmit maat ovat olleet Yhdysvallat ja Intia, joissa on kuollut paljon ihmisiä vuodepaikkojen puutteen vuoksi. Palm Beachilla Floridassa julistettiin hätätila vuodepaikkojen puutteesta johtuen elokuussa 2021 [50]. Samaan aikaan Houstonissa Texasissa saatoja potilaita odotti vuodepaikkaa, ja asiantuntijoiden mukaan ihmisiä kuolee muista syistä kuin COVID-19-viruksesta johtuen, sillä COVID-19-potilaat täyttävät vuodepaikat [51]. Syyskuun 2021 alussa Alabamassa kuoli mies sydänongelmiin, kun 43 sairaalan teho-osastot olivat täynnä ja käänyttivät miehen pois [52]. Intiassa on kuollut tuntematon määrä ihmisiä vuodepaikkojen ja lisähapen puutteeseen [53; 54].

Opinnäytetyön ja sen tulosten hyödynnettävyys [23 s. 32] koskee ensisijaisesti Espoon sairaalan hoitotyön johtamista, sillä työn perusteella voidaan alkaa suunnittelemaan ennakoivan mallin lisätutkimusta. Vasta myöhemmässä vaiheessa, mahdollisen ennakoivan mallin käyttöönoton jälkeen, voi hyöty alkaa näkymään päivittäisessä hoitotyössä.

## 7.2 Luotettavuus ja eettisyys

Eettisyyteen ja luotettavuuteen liittyen tekijän motivaation on oltava riittävä, jotta työ voidaan viedä objektiivisesti loppuun asti ilman ennako-odotuksia, sillä ne voivat vääristää lopullista kuvaa. Luotettavuuden ja kattavuuden kannalta maksullisuus ja kokotekstien saatavuus eivät saisi vaikuttaa työhön, mutta tämän opinnäytetyön resurssit huomioon ottaen niitä jouduttiin käyttämään rajauksessa. Samasta syystä julkaisemattomia tai vertaisarvioimattomia töitä ei otettu mukaan, sillä hakuprosessista olisi tullut liian laaja. [23, s. 24, 26.]

Kirjallisuuskatsausta tehdessä noudatettiin tutkimuseettisen neuvottelukunnan (TENK) laatimia ohjeita hyvästä tieteellisestä käytännöstä (HTK). Rehellisyys, huolellisuus ja tarkkaavaisuus, niin tiedonhaussa kuin tulosten esittämisessäkin, on avainasia. Kaikista mukana olleista tutkimuksista on perustiedot tallennettu, ja ne on esitelty liitteissä 1 ja 2. Opinnäytetyöaiheen ja toteuttamistavan vuoksi tutkimuslupaa tai eettistä ennakoarviointia ei vaadittu. [55.]

Opinnäytetyön luotettavuutta lisäsi riittävän korkeana pysynyt motivaatio ja tahto oppia uutta. Järjestelmällinen tiedonhaku ja luotettavat lähteet olivat myös osallisena työn luotettavuuden ylläpitämiseen. Työn luotettavuutta vähensi mahdollisuus hyödyntää Levac ym. [25] laajennetun scoping-katsauksen kehyksen parannusehdotuksia, esimerkiksi monipuolista iterointia sekä kahden tutkijan käyttämistä tiedonkeruun ja -analysoinnin aikana.

## 7.3 Mitä opin

Opinnäytetyön kaikki vaiheet ja niissä läpikäytyt aiheet olivat uusia ja vaativat huomattavasti perehtymistä. Kirjallisuuskatsausta ei ollut aiemmin toteutettu, joten alussa tuli opiskella, mikä katsaustyyppi olisi sopivin rajauksen ja työmäärän kannalta. Scoping-katsauksessa käytetty järjestelmällinen tiedonhaku oli antoisimpia vaiheita. Tiedonhaun aikana opittiin teoriaa ja käytäntöjä tiedonhausta

itsessään, eri tietokantojen laajuuksista ja minkätyyppisiin aiheisiin ne keskittyvät ja miten laajaa kirjallisuuskatsausta pitäisi aikatauluttaa. Tiedonhaku ja aineiston analysointi olivat lähes kahden kuukauden työ.

Tutkimusten analysoinnin ja tulosten pohdinnan aikana opittiin, kuinka paljon erilaisia algoritmeja ja parametreja ennakoivissa malleissa voidaan käyttää. Mielenkiintoisia olivat erilaiset pisteytykset ja luokitukset, mitä tutkimuksissa käytettiin parametreina. Työn aikana käytettiin lisäksi työkaluja ja ohjelmia, kuten Microsoft Excel ja ProQuest RefWorks, joiden käytöstä ja ominaisuuksista opittiin uutta.

#### 7.4 Jatkotutkimus ja ehdotukset

Suosittelavaa olisi, että tulevaisuudessa Espoon sairaala teettäisi tarkemman katsauksen ennakoiviin malleihin, jotta voitaisiin tarkentaa tarvittavan mallin tyyppiä. Lisäksi matemaattisia malleja tulisi tarkastella laajemmin, sillä tässä työssä siihen ei ollut mahdollisuutta. Mahdollisen varautumisen mallin käyttöönottoon voisi aloittaa tutkimalla minkälaista tietoa potilaista kerätään tällä hetkellä, ja vertaamalla sitä tulostaulukkoon liitteessä 2. Tällöin voitaisiin etukäteen miettiä, onko kerättäviä tai käytettäviä parametreja tarpeeksi ja voidaanko niille antaa riittävästi painoarvoa. Mallien tai algoritmien tarkkaa toimintaa ei käsitelty tässä opinnäytetyössä, joten sellaisen luominen, käyttöönotto ja testaus vaativat lisää perehtymistä aiheeseen.

## Lähteet

- 1 Hagerlund, Tony. 20.8.2019. Väestörakenne. Kuntaliitto. Verkkoaineisto. <<https://www.kuntaliitto.fi/tietotuotteet-ja-palvelut/analyysit-ja-tietoaineistot/kuntakuvaajat/vaesto>>. Luettu 27.9.2021.
- 2 Horppu, Anna. 11.12.2020. Miksi koronalisää ei haluta maksaa sairaanhoitajille? – “Hoitajat eivät ole joulukoristeita, jotka voi nostaa tarvittaessa esiin varastosta”. Yle. Verkkoaineisto. <<https://yle.fi/aihe/artikkeli/2020/12/11/miksi-koronalisaa-ei-haluta-maksaa-sairaanhoitajille-hoitajat-eivat-ole>>. Luettu 27.9.2021.
- 3 Gronvall, Ursula. Päivitetty 25.2.2021. 70 koronahoidossa HUS:ssa – Huolestunut ylilääkäri: Eri tilanne kuin keväällä, ”sairaalat nyt täynnä muita potilaita”. Uusisuomi. Verkkoaineisto. <<https://www.uusisuomi.fi/uutiset/70-koronahoidossa-husssa-huolestunut-ylilaakari-eri-tilanne-kuin-kevaalla-sairaalat-nyt-taynna-muita-potilaita/7b3cf671-22b4-43ad-9852-ea1ba36db392>>. Luettu 27.9.2021.
- 4 Kutafina, Ekaterina; Bechtold, Istvan; Kabino, Klaus; Jonas, Stephan M. 7.3.2019. Recursive neural networks in hospital bed occupancy forecasting. Verkkoaineisto. <<https://pubmed.ncbi.nlm.nih.gov/30845940/>>. Luettu 25.8.2021.
- 5 Santos, Javier; Wysk, Richard A.; Torres, Jose M. 2006. Improving production with Lean Thinking. E-kirja. Somerset: John Wiley & Sons, Inc.
- 6 Uusi Espoon sairaala aloitti toimintansa. 7.3.2017. Länsiväylä. Verkkoaineisto. <<https://www.lansivayla.fi/paikalliset/1737020>>. Luettu 25.8.2021.
- 7 Case Espoon sairaala. Verkkoaineisto. Fidelix. <<https://www.fidelix.fi/project/espoo-sairaala/>>. Luettu 5.4.2021.
- 8 Espoon sairaalan osastot. Espoon kaupunki. Verkkoaineisto. <<https://www.espoo.fi/fi/terveys/espoo-sairaala>>. Luettu 30.9.2021.
- 9 Potilaana Espoon sairaalassa. Espoon sairaala. Verkkoaineisto. <<https://www.espoo.fi/fi/terveys/espoo-sairaala/potilaana-espoo-sairaalassa>>. Luettu 30.9.2021.
- 10 Ortopedinen kuntoutusosasto. Espoon sairaala. Verkkoaineisto. <<https://www.espoo.fi/fi/toimipisteet/64688>>. Luettu 30.9.2021.
- 11 Kuntoutus Espoon sairaalassa. Espoon sairaala. Verkkoaineisto. <<https://www.espoo.fi/fi/terveys/espoo-sairaala/kuntoutus-espoo-sairaalassa>>. Luettu 30.9.2021.

- 12 Autti-Rämö, Ilona; Komulainen, Jorma. 2013. Kuntoutus perustuu tietoon - kuntoutumisen mahdollistavat asenteet, prosessit, osaaminen ja yksilöllisyys. LÄÄKETIETEELLINEN AIKAKAUSKIRJA DUODECIM. Verkkoaineisto. <<https://www.duodecimlehti.fi/duo10857>>. Luettu 29.9.2021.
- 13 Järvisalo, Mari. 2021. Laatu päällikkö. Espoon sairaala. Keskustelu 25.3.2021.
- 14 Daghistani, Tahani A.; Elshawi, Radwa; Sakr, Sherif; Ahmed, Amjad M.; Al-Thwayee, Abdullah; Al-Mallah, Mouaz H. 2019. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. International Journal of Cardiology. Verkkoaineisto. <<https://doi.org/10.1016/j.ijcard.2019.01.046>>. Luettu 19.8.2021.
- 15 Airila, Marika. Mitä on Lean? Leanisti kohti yhä sujuvampaa työtä. Verkkoaineisto. Talentree. <<https://talentree.fi/konsultointi/mita-on-lean/>>. Luettu 7.6.2021.
- 16 Suneja, Aneesh; Suneja, Carolyn. 27.4.2017. Lean ja terveydenhuolto. Verkkoaineisto. Duodecim oppiportti. <<https://www.oppoportti.fi/op/Int01204/do#s8>>. Luettu 14.5.2021.
- 17 Skhmot, Nawras. 5.8.2017. What is Lean? Verkkoaineisto. <<https://theleanway.net/what-is-lean>>. Luettu 22.5.2021.
- 18 Lean thinking and practice. Lean Enterprise Institute. Verkkoaineisto. <<https://www.lean.org/lexicon-terms/lean-thinking-and-practice/>>. Luettu 22.5.2021.
- 19 What is Lean Healthcare? 27.4.2018. Verkkoaineisto. NEJM Catalyst. <<https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0193>>. Luettu 10.5.2021.
- 20 Pakdil, Fatma. 2020. Six Sigma for Students. A Problem-solving Methodology. E-kirja. Springer Nature Switzerland AG.
- 21 Caetano, Nuno; Cortez, Paulo; Laureano, Raul M.S. 2015. Using Data Mining for Prediction of Hospital Length of Stay: An Application of the CRISP-DM Methodology. ICEIS 2014. Verkkoaineisto. <[https://link.springer.com/chapter/10.1007/978-3-319-22348-3\\_9](https://link.springer.com/chapter/10.1007/978-3-319-22348-3_9)>. Luettu 25.9.2021.
- 22 Tuomi, Sirpa; Latvala, Eila. Kirjallisuuskatsaukset. Opinnäytetyön ohjaajan käsikirja. Verkkoaineisto. JAMK. <<https://oppimateriaalit.jamk.fi/yamk-kasikirja/kirjallisuuskatsaukset/>>. Luettu 8.6.2021.
- 23 Stolt, Minna; Axelin, Anna; Suhonen, Riitta (toim.) 2016. Kirjallisuuskatsaus hoitotieteessä. Turun Yliopisto.

- 24 Arksey, Hillary; O'Malley, Lisa. 2005. Scoping studies: towards a methodological framework. Verkkoaineisto. International Journal of Social Research Methodology. <<https://eprints.whiterose.ac.uk/1618/1/Scopingstudies.pdf>>.
- 25 Levac, Danielle; Colquhoun, Heather; O'Brien, K.K. 2010. Scoping studies: advancing the methodology. Verkkoaineisto. Implementation Science. <<https://link.springer.com/article/10.1186/1748-5908-5-69>>. Luettu 19.7.2021.
- 26 Aveyard, Helen. 2010. Doing A Literature Review In Health And Social Care: A Practical Guide. E-kirja. McGraw-Hill Education.
- 27 Peters, Micah D.J.; Godfrey, Christina M.; Khalil, Hanan; McInerney, Patricia; Parker, Deborah; Soares, Cassia Baldini. 2015. Guidance for conducting systematic scoping reviews. Verkkoaineisto. JBI. <[https://journals.lww.com/ijebh/Fulltext/2015/09000/Guidance\\_for\\_conducting\\_systematic\\_scoping\\_reviews.5.aspx?bid=AMCampaignWKHJ](https://journals.lww.com/ijebh/Fulltext/2015/09000/Guidance_for_conducting_systematic_scoping_reviews.5.aspx?bid=AMCampaignWKHJ)>. Luettu 16.7.2021.
- 28 Munn, Z.; Peters, M.D.J.; Stern, Cindy; Tufanaru, Catalin; McArthur, Alexa; Aromataris, Edoardo. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. Verkkoaineisto. BMC Medical Research Methodology. <<https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-018-0611-x>>. Luettu 16.7.2021.
- 29 Wright, Rick; Brand, Richard; Warren, Dunn; Spindler, Kurt. 2007. How to Write a Systematic Review, Clinical Orthopaedics and Related Research. Verkkoaineisto. <<https://doi.org/10.1097/BLO.0b013e31802c9098>>. Luettu 16.7.2021.
- 30 Kastner, Monika; Tricco, Andrea.C.; Soobiah, Charlene. et al. 2012. What is the most appropriate knowledge synthesis method to conduct a review? Protocol for a scoping review. Verkkoaineisto. BMC Medical Research Methodology. <<https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-12-114#citeas>>. Luettu 16.7.2021.
- 31 Mäkelä, Marjukka; Punkari, Kerttuli. 2017. Käsitteitä. Verkkoaineisto. Terveysportti. <[https://www.terveysportti.fi/dtk/hta/avaa?p\\_artikkeli=hta00002](https://www.terveysportti.fi/dtk/hta/avaa?p_artikkeli=hta00002)>. Luettu 14.6.2021.
- 32 Malmivaara, Antti; Komulainen, Jorma. 2014. Luotettavaa vaikuttavuustietoa järjestelmällisistä katsauksista. Verkkoaineisto. Duodecim. <<https://www.duodecimlehti.fi/duo11791>>. Luettu 20.7.2021.

- 33 Isojärvi, Jaana. 2017. Kirjallisuushaku. Verkkoaineisto. Terveysportti. <[https://www.terveysportti.fi/dtk/hta/avaa?p\\_artikkeli=hta00008](https://www.terveysportti.fi/dtk/hta/avaa?p_artikkeli=hta00008)>. Luettu 14.6.2021.
- 34 Vuori, Jaana. Laadullinen sisällönanalyysi. Teoksessa Jaana Vuori (toim.) Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. <<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/laadullinen-sisallanalyysi/>>. Luettu 22.9.2021.
- 35 Juhila, Kirsi. Koodaaminen. Teoksessa Jaana Vuori (toim.) Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. <<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/koodaaminen/>>. Luettu 22.9.2021.
- 36 ACS National Surgical Quality Improvement Program. American College of Surgeons. Verkkoaineisto. <<https://www.facs.org/quality-programs/acs-nsqip>>. Luettu 24.9.2021.
- 37 Koneoppiminen. Digitalisoinnin opas. Itewiki. Verkkoaineisto. <<https://www.itewiki.fi/opas/koneoppiminen/>>. Luettu 23.9.2021.
- 38 ASA Physical Status Classification System. Päivitetty 13.12.2020. American Society of Anesthesiologists. Verkkoaineisto. <<https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system>>. Luettu 24.9.2021.
- 39 Kabir, S.; Farrokhvar, L. 2019. Non-Linear Feature Selection for Prediction of Hospital Length of Stay. Verkkoaineisto. <<https://doi.org/10.1109/ICMLA.2019.00162>>. Luettu 14.9.2021.
- 40 Stecker, M. M., Stecker, M., & Falotico, J. 2017. Predictive model of length of stay and discharge destination in neuroscience admissions. Surgical Neurology International. Verkkoaineisto. <<https://doi.org/10.4103/2152-7806.199558>>. Luettu 10.9.2021.
- 41 Abd-Elrazek, Merhan A.; Eltahawi, Ahmed A.; Abd Elaziz, Mohamed H.; Abd-Elwhab, Mohamed N. 2021. Predicting length of stay in hospitals intensive care unit using general admission features. Ain Shams Engineering Journal. Verkkoaineisto. <<https://doi.org/10.1016/j.asej.2021.02.018>>. Luettu 5.9.2021.
- 42 Barnes, Sean; Hamrock, Eric; Toerper, Matthew; Siddiqui, Sauleh; Levin, Scott. 2016. Real-time prediction of inpatient length of stay for discharge prioritization. Journal of the American Medical Informatics Association. Verkkoaineisto. <<https://doi.org/10.1093/jamia/ocv106>>. Luettu 4.9.2021.

- 43 Harerimana, Gaspard; Kim, Jong Wook; Jang, Beakcheol. 2021. A deep attention model to forecast the Length Of Stay and the in-hospital mortality right on admission from ICD codes and demographic data. *Journal of Bio-medical Informatics*. Verkkoaineisto. <<https://doi.org/10.1016/j.jbi.2021.103778>>. Luettu 9.9.2021.
- 44 Cereda, Emanuele; Klersy, Catherine; Pedrolli, Carlo; Cameletti, Barbara; Bonardi, Chiara; Quarleri, Lara; Cappello, Silvia; Bonoldi, Alberto; Bonadeo, Elisa; Caccialanza, Riccardo. 2015. The Geriatric Nutritional Risk Index predicts hospital length of stay and in-hospital weight loss in elderly patients. *Clinical Nutrition*. Verkkoaineisto. <<https://doi.org/10.1016/j.clnu.2014.01.017>>. Luettu 8.9.2021.
- 45 Pettersson, Tom. 2001. Hypoalbuminemia ja sen kliininen merkitys. LÄÄKETIETEELLINEN AIKAKAUSKIRJA DUODECIM. Verkkoaineisto. <<https://www.duodecimlehti.fi/duo92483>>. Luettu 28.9.2021.
- 46 Moreault, Olivier; Lacasse, Yves; Bussi eres, Jean S. 2017. Calculating Ideal Body Weight: Keep It Simple. *Anesthesiology*. Verkkoaineisto. <<https://doi.org/10.1097/ALN.0000000000001687>>. Luettu 28.9.2021.
- 47 Verburg, IW.; Atashi, A.; Eslami, S.; Holman, R.; Abu-Hanna A., de Jonge, E.; Peek, N; de Keizer, NF. 2017. Which Models Can I Use to Predict Adult ICU Length of Stay? A Systematic Review. *Crit Care Med*. 45(2):e222-e231. <<https://doi.org/10.1097/ccm.0000000000002054>>. Luettu 25.9.2021.
- 48 Henke, Nicolaus; Puri, Ankur; Saleh, Tamim. 2020. Accelerating analytics to navigate COVID-19 and the next normal. *McKinsey Analytics*. E-kirja.
- 49 Gauffin, Jarno. 2020. COVID-19-MALLINNUS SOTILASLÄÄKETIETEEN KESKUKSESSA. *Sotilasl aaketieteen aikakauslehti 2/2020*. Verkkoaineisto. <[https://puolustusvoimat.fi/documents/1948673/9261842/Sotilaslaaketieteen\\_aikakauslehti\\_2\\_2020.pdf/b583b4d8-a59f-b3dc-6f8b-7630e8e6aab8/Sotilaslaaketieteen\\_aikakauslehti\\_2\\_2020.pdf?t=1608291717343](https://puolustusvoimat.fi/documents/1948673/9261842/Sotilaslaaketieteen_aikakauslehti_2_2020.pdf/b583b4d8-a59f-b3dc-6f8b-7630e8e6aab8/Sotilaslaaketieteen_aikakauslehti_2_2020.pdf?t=1608291717343)>. Luettu 16.9.2021.
- 50 Papaycik, Matt. P aivitetty 18.8.2021. Palm Beach County declares state of emergency over COVID-19 pandemic, hospital bed shortage. Verkkoaineisto. WPTV. <<https://www.wptv.com/coronavirus/palm-beach-county-declares-state-of-emergency-over-covid-19-pandemic-hospital-bed-shortage>>. Luettu 20.9.2021.
- 51 Zubowski, Courtney. 20.8.2021. 'I've never seen anything like this': ER doctor says hundreds of patients are waiting to be admitted into hospitals, but there are no beds. Verkkoaineisto. Click2Houston. <<https://www.click2houston.com/news/local/2021/08/20/ive-never-seen->



anything-like-this-houston-doctor-says-hundreds-of-patients-are-waiting-to-be-admitted-into-hospitals-but-there-are-no-beds/>. Luettu 20.9.2021.

- 52 Hitson, Hadley. 10.9.2021. Alabama man dies of cardiac event after 43 hospitals with full ICUs turned him away. Verkkoaineisto. UsaToday. <<https://eu.usatoday.com/story/news/health/2021/09/10/cullman-al-man-dies-hospitals-full-icus-turn-away-covid/8281712002/>>. Luettu 20.9.2021.
- 53 Covid-19: India outrage over 'no oxygen shortage death data' claim. 21.7.2021. Verkkoaineisto. BBC. <<https://www.bbc.com/news/world-asia-india-57911638>>. Luettu 20.9.2021.
- 54 Pandey, Vikas. 18.5.2021. India coronavirus: Tracing a family's journey to save a life. Verkkoaineisto. BBC. <<https://www.bbc.com/news/world-asia-india-57111161>>. Luettu 20.9.2021.
- 55 Hyvä tieteellinen käytäntö. Tutkimuseettinen neuvottelukunta (TENK). Verkkoaineisto. <<https://tenk.fi/fi/tiedevilppi/hyva-tieteellinen-kaytanto-htk>>. Luettu 20.9.2021.

**Liite 1: Kokotekstivalinnan artikkelit**

A. WARD, A. MANN, J. VALLON, G. ESCOBAR, N. BAMBOS and A. SCHULER, 2021. Operationally-Informed Hospital-Wide Discharge Prediction Using Machine Learning, - 2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM) 2021, pp. 1–6.

ABD-ELRAZEK, M.A., ELTAHAWI, A.A., ABD ELAZIZ, M.H. and ABD-ELWHAB, M.N., 2021. Predicting length of stay in hospitals intensive care unit using general admission features. *Ain Shams Engineering Journal*-

ALLAUDEEN, N., SCHNIPPER, J.L., ORAV, E.J., WACHTER, R.M. and VIDYARTHI, A.R., 2011. Inability of providers to predict unplanned readmissions. *Journal of General Internal Medicine*, 26(7), pp. 771–776.

ALMEIDA, JOÃO FLÁVIO DE FREITAS, CONCEIÇÃO, S.V., PINTO, L.R., HORTA, C.J.G., MAGALHÃES, V.S. and CAMPOS, FRANCISCO CARLOS CARDOSO DE, 2021. Estimating Brazilian states' demands for intensive care unit and clinical hospital beds during the COVID-19 pandemic: development of a predictive model. *Sao Paulo Medical Journal = Revista Paulista De Medicina*, 139(2), pp. 178–185.

ARKUN, A., BRIGGS, W.M., PATEL, S., DATILLO, P.A., BOVE, J. and BIRKHAHN, R.H., 2010. Emergency department crowding: factors influencing flow. *The Western Journal of Emergency Medicine*, 11(1), pp. 10–15.

BAAS, S., DIJKSTRA, S., BRAAKSMA, A., VAN ROOIJ, P., SNIJDERS, F.J., TIEMESSEN, L. and BOUCHERIE, R.J., 2021. Real-time forecasting of COVID-19 bed occupancy in wards and Intensive Care Units. *Health Care Management Science*, 24(2), pp. 402–419.

BACCHI, S., GLUCK, S., TAN, Y., CHIM, I., CHENG, J., GILBERT, T., MENON, D.K., JANNES, J., KLEINIG, T. and KOBLAR, S., 2020. Prediction of general medical admission length of stay with natural language processing and deep learning: a pilot study. *Internal & Emergency Medicine*, 15(6), pp. 989–995.

BARNES, S., HAMROCK, E., TOERPER, M., SIDDIQUI, S. and LEVIN, S., 2016. Real-time prediction of inpatient length of stay for discharge prioritization. *Journal of the American Medical Informatics Association: JAMIA*, 23(e1), pp. e2-e10.

BELCIUG, S. and GORUNESCU, F., 2015. Improving hospital bed occupancy and resource utilization through queuing modeling and evolutionary computation. *Journal of Biomedical Informatics*, 53, pp. 261–269.

- BERRIOS, R.A., 2019. Forecasting Patient Discharge Before Noon: A Comparison Between Holt's and Box-Jenkins' Models. *Quality Management in Health Care*, 28(4), pp. 237–244.
- BOUCKAERT, N., VAN DEN HEEDE, K. and VAN DE VOORDE, C., 2018. Improving the forecasting of hospital services: A comparison between projections and actual utilization of hospital services. *Health Policy*, 122(7), pp. 728–736.
- BRANDI, S., TROSTER, E.J. and CUNHA, MARIANA LUCAS DA ROCHA, 2020. Length of stay in pediatric intensive care unit: prediction model. *Einstein (Sao Paulo, Brazil)*, 18, pp. eAO5476.
- BUCCI, S., DE BELVIS, A.G., MARVENTANO, S., DE LEVA, A.C., TANZARIELLO, M., SPECCHIA, M.L., RICCIARDI, W. and FRANCESCHI, F., 2016. Emergency Department crowding and hospital bed shortage: is Lean a smart answer? A systematic review. *European Review for Medical and Pharmacological Sciences*, 20(20), pp. 4209–4219.
- BULLINGHAM, R., WILKINSON, C., HELENA, P., DEAN, N. and CLAIRE, S., 2017. Predicting Length of Stay, Functional Outcomes and Discharge Destinations in Patients Presenting to Hospital with Hip Fractures. *International Journal of Surgery*, 47, pp. S83.
- C. GALLESE, E. FALLETTI, M. S. NOBILE, L. FERRARIO, F. SCHETTINI and E. FOGLIA, 2020. Preventing litigation with a predictive model of COVID-19 ICUs occupancy, - 2020 IEEE International Conference on Big Data (Big Data) 2020, pp. 2111–2116.
- CEREDA, E., KLERSY, C., PEDROLLI, C., CAMELETTI, B., BONARDI, C., QUARLERI, L., CAPPELLO, S., BONOLDI, A., BONADEO, E. and CACCIALANZA, R., 2015. The Geriatric Nutritional Risk Index predicts hospital length of stay and in-hospital weight loss in elderly patients. *Clinical Nutrition*, 34(1), pp. 74–78.
- CHIN, V., SAMIA, N.I., MARCHANT, R., ROSEN, O., IOANNIDIS, J.P.A., TANNER, M.A. and CRIPPS, S., 2020. A case study in model failure? COVID-19 daily deaths and ICU bed utilisation predictions in New York state. *European Journal of Epidemiology*, 35(8), pp. 733–742.
- DAGHISTANI, T.A., ELSHAWI, R., SAKR, S., AHMED, A.M., AL-THWAYEE, A. and AL-MALLAH, M.H., 2019. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. *International journal of cardiology*, 288, pp. 140–147.
- DAUNCEY, S.J., KELLY, P.A., BAYKOV, D., SKELDON, A.C. and WHYTE, M.B., 2021. Rhythmicity of patient flow in an acute medical unit: relationship to

hospital occupancy, seven-day working, and the effect of COVID-19. QJM: monthly journal of the Association of Physicians.

DAVIS, S. and FARD, N., 2020. Theoretical bounds and approximation of the probability mass function of future hospital bed demand. *Health Care Management Science*, 23(1), pp. 20–33.

DEVAPRIYA, P., STRÖMBLAD, C., BAILEY, M., FRAZIER, S., BULGER, J., KEMBERLING, S. and WOOD, K., 2015. StratBAM: A Discrete-Event Simulation Model to Support Strategic Hospital Bed Capacity Decisions. *Journal of medical systems*, 39(10), pp. 1–13.

DONKER, T., BÜRKIN, F.M., WOLKEWITZ, M., HAVERKAMP, C., CHRISTOFFEL, D., KAPPERT, O., HAMMER, T., BUSCH, H., BIEVER, P., KALBHENN, J., BÜRKLE, H., KERN, W.V., WENZ, F. and GRUNDMANN, H., 2021. Navigating hospitals safely through the COVID-19 epidemic tide: Predicting case load for adjusting bed capacity. *Infection Control and Hospital Epidemiology*, 42(6), pp. 653–658.

DURAND, A., D'AMOURS, L., GIROUX, A., PELLETIER, M., LEBLOND, J. and RICHARDS, C.L., 2020. Benchmarking length of stay for inpatient stroke rehabilitation without adversely affecting functional outcomes. *Journal of Rehabilitation Medicine*, 52(10), pp. jrm00113.

E OLIVEIRA, B. R. P., DE VASCONCELOS, J.A., ALMEIDA, J.F.F. and PINTO, L.R., 2020. A Simulation-Optimisation approach for hospital beds allocation. *International journal of medical informatics*, 141, pp. 104174.

ELTOUKHY, A.E.E., SHABAN, I.A., CHAN, F.T.S. and ABDEL-AAL, M.A.M., 2020. Data Analytics for Predicting COVID-19 Cases in Top Affected Countries: Observations and Recommendations. *International Journal of Environmental Research and Public Health*, 17(19).

GABRIEL, G.T., CAMPOS, A.T., MAGACHO, A.D.L., SEGISMONDI, L.C., VILELA, F.F., DE QUEIROZ, J.A. and MONTEVECHI, J.A.B., 2020. Lean thinking by integrating with discrete event simulation and design of experiments: an emergency department expansion. *PeerJ. Computer Science*, 6, pp. e284.

GAUFFIN, J., *Sotilaslääketieteen aikakauslehti*, Teemanumero COVID-19. Sotilaslääketieteen keskus.

GOPAKUMAR, S., TRAN, T., LUO, W., PHUNG, D. and VENKATESH, S., 2016. Forecasting Daily Patient Outflow From a Ward Having No Real-Time Clinical Data. *JMIR medical informatics*, 4(3), pp. e25.

H. XU, W. WU, S. NEMATI and H. ZHA, 2017. Patient Flow Prediction via Discriminative Learning of Mutually-Correcting Processes.

HARERIMANA, G., KIM, J.W. and JANG, B., 2021. A deep attention model to forecast the Length Of Stay and the in-hospital mortality right on admission from ICD codes and demographic data. *Journal of Biomedical Informatics*, 118, pp. 103778.

HECHT, G., SLEE, C.A., GOODELL, P.B., TAYLOR, S.L. and WOLINSKY, P.R., 2019. Predictive Modeling for Geriatric Hip Fracture Patients: Early Surgery and Delirium Have the Largest Influence on Length of Stay. *The Journal of the American Academy of Orthopaedic Surgeons*, 27(6), pp. e293-e300.

HOLM, L.B., LURÅS, H. and DAHL, F.A., 2013. Improving hospital bed utilisation through simulation and optimisation: With application to a 40 % increase in patient volume in a Norwegian general hospital. *International journal of medical informatics*, 82(2), pp. 80–89.

HU, Y., DONG, J., PERRY, O., CYRUS, R.M., GRAVENOR, S. and SCHMIDT, M.J., 2021. Use of a Novel Patient-Flow Model to Optimize Hospital Bed Capacity for Medical Patients. *The Joint Commission Journal on Quality and Patient Safety*, 47(6), pp. 354–363.

HWANG, U., MCCARTHY, M.L., ARONSKY, D., ASPLIN, B., CRANE, P.W., CRAVEN, C.K., EPSTEIN, S.K., FEE, C., HANDEL, D.A., PINES, J.M., RATHLEV, N.K., SCHAFERMEYER, R.W., ZWEMER, F.L. and BERNSTEIN, S.L., 2011. Measures of crowding in the emergency department: a systematic review. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, 18(5), pp. 527–538.

J. CHANG JUNIOR and D. CARUSO, 2013. Development of methodology for forecasting hospital occupancy by econometric method, - 2013 Proceedings of PICMET '13: Technology Management in the IT-Driven Services (PICMET) 2013, pp. 2517–2525.

J. LE LAY, V. AUGUSTO, X. XIE, E. ALFONSO-LIZARAZO, B. BONGUE, T. CELARIER, R. GONTHIER and M. MASMOUDI, 2020. Impact of COVID-19 Epidemics on Bed Requirements in a Healthcare Center Using Data-Driven Discrete-Event Simulation, - 2020 Winter Simulation Conference (WSC) 2020, pp. 771–781.

J. M. P. GUTIÉRREZ, M. -Á. SICILIA, S. SANCHEZ-ALONSO and E. GARCÍA-BARRIOCANAL, 2021. Predicting Length of Stay Across Hospital Departments.

JILANI, T., HOUSLEY, G., FIGUEREDO, G., TANG, P., HATTON, J. and SHAW, D., 2019. Short and Long term predictions of Hospital emergency department attendances. *International Journal of Medical Informatics*, 129, pp. 167–174.

KHO, A.N., IVORY, C., OSMUNDSON, S. and MALIN, B.A., 2017. Predicting Length of Stay for Obstetric Patients via Electronic Medical Records...16 World Congress of Medical and Health Informatics: Precision Healthcare Through Informatics (MedInfo2017) was held in Hangzhou, China from August 21st to 25th, 2017. *Studies in Health Technology & Informatics*, 245, pp. 1019–1023.

KNOLL, O.M., LAKOMKIN, N., SHEN, M.S., ADEBAYO, M., KOTHARI, P., DODD, A.C., ATTUM, B., LEE, N., CHONA, D. and SETHI, M.K., 2019. A predictive model for increased hospital length of stay following geriatric hip fracture. *Journal of Clinical Orthopaedics and Trauma*, 10(Suppl 1), pp. S84-S87.

KUTAFINA, E., BECHTOLD, I., KABINO, K. and JONAS, S.M., 2019. Recursive neural networks in hospital bed occupancy forecasting. *BMC medical informatics and decision making*, 19(1), pp. 39.

L. GARG, S. I. MCCLEAN, M. BARTON, B. J. MEENAN and K. FULLERTON, 2012. *Intelligent Patient Management and Resource Planning for Complex, Heterogeneous, and Stochastic Healthcare Systems*.

L. GARG, S. MCCLEAN, M. BARTON, B. MEENAN and K. FULLERTON, 2010. Forecasting hospital bed requirements and cost of care using phase type survival trees, - 2010 5th IEEE International Conference Intelligent Systems 2010, pp. 185–190.

L. LUO, X. XU, J. LI and W. SHEN, 2017. Short-term forecasting of hospital discharge volume based on time series analysis, - 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom) 2017, pp. 1–6.

LAPPALAINEN, K. and KIVINEN, P., 2017. Tietämyksenhallinnan avulla ennustamisesta aktiiviseen ennakointiin. *Finnish Journal of eHealth and eWelfare*, 9(1), pp. 31–36.

LECLERC, Q.J., FULLER, N.M., KEOGH, R.H., DIAZ-ORDAZ, K., SEKULA, R., SEMPLE, M.G., BAILLIE, J.K., OPENSHAW, P.J.M., CARSON, G., ALEX, B., BACH, B., BARCLAY, W.S., BOGAERT, D., CHAND, M., COOKE, G.S., DOCHERTY, A.B., DUNNING, J., DA SILVA FILIPE, A. and FLETCHER, T., 2021. Importance of patient bed pathways and length of stay differences in predicting COVID-19 hospital bed occupancy in England. *BMC Health Services Research*, 21(1), pp. 1–15.

LEVIN, SCOTT R.; HARLEY, ERIC T.; FACKLER, JAMES C.; LEHMANN, CHRISTOPH U.; Custer, Jason W.; France, Daniel; Zeger, Scott L. Real-time forecasting of pediatric intensive care unit length of stay using computerized provider orders. 2012. *Critical Care Medicine*, 40(11), pp. 3058–3064.

LISK, R., UDDIN, M., PARBHOO, A., YEONG, K., FLUCK, D., SHARMA, P., LEAN, M.E.J. and HAN, T.S., 2019. Predictive model of length of stay in hospital among older patients. *Aging Clinical & Experimental Research*, 31(7), pp. 993–999.

MAHARLOU, H., NIAKAN KALHORI, S.R., SHAHBAZI, S. and RAVANGARD, R., 2018. Predicting Length of Stay in Intensive Care Units after Cardiac Surgery: Comparison of Artificial Neural Networks and Adaptive Neuro-fuzzy System. *Healthcare Informatics Research*, 24(2), pp. 109–117.

MIXON, A.S., GOGGINS, K., BELL, S.P., VASILEVSKIS, E.E., NWOSU, S., SCHILDCROUT, J.S. and KRIPALANI, S., 2016. Preparedness for hospital discharge and prediction of readmission. *Journal of Hospital Medicine*, 11(9), pp. 603–609.

NOVATI, R., PAPALIA, R., PEANO, L., GORRAZ, A., ARTUSO, L., CANTA, M.G., DEL VESCOVO, G. and GALOTTO, C., 2017. Effectiveness of an hospital bed management model: results of four years of follow-up. *Annali Di Igiene: Medicina Preventiva E Di Comunita*, 29(3), pp. 189–196.

P. LANDA, M. SONNESSA, E. TÀN FANI and A. TESTI, 2014. A Discrete Event Simulation model to support bed management, - 2014 4th International Conference On Simulation And Modeling Methodologies, Technologies And Applications (SIMULTECH) 2014, pp. 901–912.

PAGEL, C., BANKS, V., POPE, C., WHITMORE, P., BROWN, K., GOLDMAN, A. and UTLEY, M., 2017. Development, implementation and evaluation of a tool for forecasting short term demand for beds in an intensive care unit. *Operations Research for Health Care*, 15, pp. 19–31.

PECK, J.S., BENNEYAN, J.C., NIGHTINGALE, D.J. and GAEHDE, S.A., 2012. Predicting emergency department inpatient admissions to improve same-day patient flow. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, 19(9), pp. 1045.

PECORARO, F., LUZI, D. and CLEMENTE, F., 2021. The efficiency in the ordinary hospital bed management: A comparative analysis in four European countries before the COVID-19 outbreak. *PLoS One*, 16(3), pp. e0248867.

PEREZ-CONCHA, O., COIERA, E., MARTIN-SANCHEZ, F., DAY, R., ROFFE, D., GALLEGO, B. and CAI, X., 2016. Real-time prediction of mortality, readmission, and length of stay using electronic health record data. *Journal of the American Medical Informatics Association*, 23(3), pp. 553–561.

R. BARES, J. GRIFFITHS, V. KNIGHT, J. WILLIAMS, K. BABOOLAL and A. NELSON, 2012. Simulating Bed Capacity: Evaluating the Impact of Healthcare Service Transfers, - 2012 UKSim 14th International Conference on Computer Modelling and Simulation 2012, pp. 358–362.

RÖMMELE, C., NEIDEL, T., HEINS, J., HEIDER, S., OTTEN, V., EBIGBO, A., WEBER, T., MÜLLER, M., SPRING, O., BRAUN, G., WITTMANN, M., SCHOENFELDER, J., HELLER, A.R., MESSMANN, H. and BRUNNER, J.O., 2020. [Bed capacity management in times of the COVID-19 pandemic: A simulation-based prognosis of normal and intensive care beds using the descriptive data of the University Hospital Augsburg]. *Der Anaesthetist*, 69(10), pp. 717–725.

RUYSSINCK, J., VAN DER HERTEN, J., HOUTHOOFT, R., ONGENAE, F., COUCKUYT, I., GADEYNE, B., COLPAERT, K., DECRUYENAERE, J., DE TURCK, F. and DHAENE, T., 2016. Random Survival Forests for Predicting the Bed Occupancy in the Intensive Care Unit. *Computational and Mathematical Methods in Medicine*, 2016, pp. 7087053.

S. D. SOUSA, C. S. RODRIGUES and E. P. NUNES, 2016. On the estimation of hospital beds occupancy after hip surgery, - 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) 2016, pp. 691–695.

S. GOPAKUMAR, T. TRAN, W. LUO, D. PHUNG and S. VENKATESH, 2016. Forecasting Patient Outflow from Wards having No Real-Time Clinical Data, - 2016 IEEE International Conference on Healthcare Informatics (ICHI) 2016, pp. 177–183.

S. KABIR and L. FARROKHVAR, 2019. Non-Linear Feature Selection for Prediction of Hospital Length of Stay, - 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA) 2019, pp. 945–950.

S. NAGPAL, V. A. ATHAVALE, A. K. SAINI and R. SHARMA, 2020. Indian Health Care System is Ready to Fight Against COVID-19 A Machine Learning Tool for Forecast the Number of Beds, - 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) 2020, pp. 61–65.

SCHMIDT, R., GEISLER, S. and SPRECKELSEN, C., 2013. Decision support for hospital bed management using adaptable individual length of stay estimations and shared resources. *BMC Medical Informatics & Decision Making*, 13(1), pp. 3–3.



SEEMATTER-BAGNOUD, L., FUSTINONI, S., DUNG, D.H., SANTOS-EG-GIMANN, B., KOEHN, V., BIZE, R., OETTLI, A. and WASSERFALLEN, J.-., 2015. Comparison of different methods to forecast hospital bed needs. *European Geriatric Medicine*, 6(3), pp. 262–266.

SHAIKH, S.A., ROBINSON, R.D., CHEETI, R., RATH, S., COWDEN, C.D., ROSINIA, F., ZENAROSA, N.R. and WANG, H., 2018. Risks predicting prolonged hospital discharge boarding in a regional acute care hospital. *BMC health services research*, 18(1), pp. 59.

STECKER, M.M., STECKER, M. and FALOTICO, J., 2017. Predictive model of length of stay and discharge destination in neuroscience admissions. *Surgical Neurology International*, 8, pp. 17.

T. LATRUWE, M. V. D. WEE, J. DEVRIESE, P. VANLEENHOVE, S. VERBRUGGE and D. COLLE, 2021. A Model for Inpatient Ward Simulation and Forecasting, - 2021 7th International Conference on Information Management (ICIM) 2021, pp. 71–76.

TAKEUCHI, S. and KURODA, Y., 2010. [Predicting spread of new pandemic swine-origin influenza A (H1N1) in local mid-size city: evaluation of hospital bed shortage and effectiveness of vaccination]. *Nihon Eiseigaku Zasshi. Japanese Journal of Hygiene*, 65(1), pp. 48–52.

TEOW, K.L., TAN, K.B., PHUA, H.P. and ZHU, Z., 2018. Applying Gravity model to predict demand of public hospital beds. *Operations Research for Health Care*, 17, pp. 65–70.

TSAI, M., ABERNETHY, N., MITCHELL, S., SUN, J., TSAI, K. and PAYNE, T., 2018. 16 Forecasting Medical Patient Length of Stay at Presentation in an Emergency Department Using Machine Learning. *Annals of Emergency Medicine*, 72(4, Supplement), pp. S8.

V. TIWARI and W. S. SANDBERG, 2016. Perioperative bed capacity planning guided by theory of constraints, - 2016 Winter Simulation Conference (WSC) 2016, pp. 1894–1903.

VISWESWARAN, S., ANGUS, D.C., HSIEH, M., WEISSFELD, L., YEALY, D. and COOPER, G.F., 2010. Learning patient-specific predictive models from clinical data. *Journal of Biomedical Informatics*, 43(5), pp. 669–685.

WARD, T.L., RAYBOULD, S.J., MONDAL, A., LAMBERT, J. and PATEL, B., 2021. Predicting the length of stay at admission for emergency general surgery patients a cohort study. *Annals of Medicine and Surgery*, 62, pp. 127–130.

WARDE, P.R., PATEL, S., FERREIRA, T., GERSHENGORN, H., BHATIA, M.C., PAREKH, D., MANNI, K. and SHUKLA, B., 2021. Linking prediction models to government ordinances to support hospital operations during the COVID-19 pandemic. *BMJ health & care informatics*, 28(1).

WOOD, R.M., 2021. Nowcasting for improved management of COVID-19 acute bed capacity. *British journal of healthcare management*, 27(2), pp. 1–3.

YUE-JIAO GONG, J. ZHANG and ZHUN FAN, 2010. A multi-objective comprehensive learning particle swarm optimization with a binary search-based representation scheme for bed allocation problem in general hospital, - 2010 IEEE International Conference on Systems, Man and Cybernetics 2010, pp. 1083–1088.

## Liite 2: Tulostaulukko analysoiduista artikkeleista

Otsikko	Real-time forecasting of pediatric intensive care unit length of stay using computerized provider orders	Operationally-Informed Hospital-Wide Discharge Prediction Using Machine Learning
Tekijät	Levin, Scott R.; Harley, Eric T.; Fackler, James C.; Lehmann, Christoph U.; Custer, Jason W.; France, Daniel; Zeger, Scott L.	A. Ward, A. Mann, J. Vallon, G. Escobar, N. Bambos and A. Schuler,
Julkaisu vuosi	2012	2021
Tutkimustyyppi	Retrospective cohort study	
Otannan suuruus (henkilöä/vastaanottoa)	2178	243696
Otannan kesto (kk)	16	25
Mallin tyyppi	joint discrete-time logistic regression model	machine learning
Algoritmi(t)		logistic regression (LR), logistic regression with a least absolute shrinkage and selection operator penalty (Lasso), random forest (RF), and gradient boosted machine (GBM)
Parametrit	Age, source, readmission, timing, medication, ventilation, activity, laboratory, diet, foreign body, Extracorporeal membrane oxygenation	age, gender, admit type (admitted from emergency department (ED), skilled nursing facility (SNF), or elsewhere), admission and hourly LAPS2 (physiologic severity of illness) and COPS2 (12-month longitudinal comorbidity) scores, patients' admission diagnoses (grouped into one of 31 different diagnosis classes, or Primary Conditions, e.g. cancer, surgical, trauma),
Ennakoinnin tarkkuus ja pituus	accurately forecasting discharge within a 12-hr window: 46% for patients within 1 day of discharge, 34% for patients within 2 days of discharge, and 27% for patients within 3 days of discharge.	best-performing models for both the 3PM and 3AM predictions is the GBM, with average precision of 0.598 and 0.419 (kotiutus seuraavan 14 tai 26 tunnin aikana)
Suosittelvatko tutkijat käyttöä	Vaatii lisätutkimusta	Kyllä

Predicting length of stay in hospitals intensive care unit using general admission features	Real-time forecasting of COVID-19 bed occupancy in wards and Intensive Care Units	Prediction of general medical admission length of stay with natural language processing and deep learning: a pilot study
Merhan A. Abd-Elrazek, Ahmed A. Eltahawi, Mohamed H. Abd Elaziz, Mohamed N. Abd-Elwhab	Stef Baas, Sander Dijkstra, Aleida Braaksma, Plom van Rooij, Fieke J. Snijders, Lars Tiemessen & Richard J. Boucherie	Bacchi, Stephen; Gluck, Samuel; Tan, Yiran; Chim, Ivana; Cheng, Joy; Gilbert, Toby; Menon, David K.; Jannes, Jim; Kleinig, Timothy and Koblar, Simon
2021	2021	2020
		pilot study
233		313
24	4	2
machine learning	mathematical model	machine learning
Neural Networks (NN), Classification Trees (CT), Tree Bagger (TB), Random Forest (RF), Fuzzy Logic (FL), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Regression Tree (RT) and Naive Bayes (NB).	Poisson Arrival Location Model (PALM), Richards' curve, Kaplan-Meier estimator, Monte-Carlo method, Levenberg-Marquardt algorithm	logistic regression, random forest models, support vector machines, and DL (deep learning), convolutional and artificial neural network (ANN), natural language processing (NLP)
(age and gender), vital signs features (blood pressure, temperature, respiratory rate, etc), basic lab features (CBC, Na, K, creatinine, ABG), chronic illness features (HTN, DM, IHD, etc), basic diagnose features (Cardiac, Chest, Stroke, etc), and Glasgow Coma Scale (GCS), LoS (taulukko tutkimuksessa, 31 parametria)	set of time stamps for patient admissions, transfers and discharges, age, gender, BMI and treatment restrictions, historic data of covid LoS's	for NLP: demographic details, the Emergency Department (ED) triage note, the note by the ED doctor, and initial investigation results
Alle 2 päivästä yli viikkoon. Tarkimmat FL 91.99 ± 7.43 % ja CT 90.04 ± 4.05 %	95% käyttöasteeseen tähdäten 1-5 päivän ennakkoinnin tarkkuus vaihteli 78-99%:n välillä teho-osastolla, ja 57-84%:n välillä vuodeosastolla	LOS yli 2 päivää, ANN tarkkuus 82%. Kotiutumispäivän tarkkuus logistic regression 76%, ANN 74%.
Kyllä	Kyllä	Vaatii lisätutkimusta

Real-time prediction of inpatient length of stay for discharge prioritization	Length of stay in pediatric intensive care unit: prediction model	The Geriatric Nutritional Risk Index predicts hospital length of stay and in-hospital weight loss in elderly patients
Barnes, Sean; Hamrock, Eric; Toerper, Matthew; Siddiqui, Sauleh; Levin, Scott	Brandi, Simone; Troster, Eduardo Juan and Cunha, Mariana Lucas da Rocha	Emanuele Cereda, Catherine Klersy, Carlo Pedrolli, Barbara Cameletti, Chiara Bonardi, Lara Quarleri, Silvia Cappello, Alberto Bonoldi, Elisa Bonadeo, Riccardo Caccialanza
2016	2020	2015
	retrospective cohort study	prospective multicentre hospital-based cohort study
9636	1815	667
26	30	
machine learning	Logistic models	Geriatric Nutritional Risk Index (GNRI)
logistic regression, classification and regression trees, tree-based ensemble learning	logistic function linking and structure of autoregressive correlation	logistic regression
patient census, day of the week, elapsed length of stay, age, gender, ethnicity, insurance, reason for visit, observation status, pending discharge location	age, sex, indication and type for admission (elective, urgency, or emergency), outcome (transference, hospital discharge (with homecare), or death), mechanical ventilation, origin (pediatric unit, emergency unit, surgical center, pediatric outpatient unit, bone marrow transplantation or external), readmission <48 hours, reason for admission (respiratory failure, sepsis, shock, post-operative, liver failure, neurology, hemodynamic monitoring, post-event monitoring, other), Paediatric Logistic Organ Dysfunction (PELOD) score, and presence of venous access.	gender, age, ward admission, main diagnosis, number of comorbidities (excluding malnutrition) and nutritional parameters (height (measured or estimated from knee-height, when this proxy indicator was more accurate), weight, body mass index (BMI; in kg/m <sup>2</sup> ) and serum albumin.).
Kyseisen päivän ennakointi RRF (regression random forest) mallilla: kotiutuminen ennen klo 2, tarkkuus 82%, kotiutuminen päivän loppuun mennessä tarkkuus 63%.	The accuracy model from 3 to 4 days was 0.71 and model greater than 4 days was 0.69. The accuracy found for 3 to 4 days (65%) and greater than 4 days (66%)	Patients with a high nutritional risk were more likely (OR = 1.89; 95%CI: 1.22–2.92) to stay longer in hospital (fourth quartile, LOS ≥ 20 days) compared to those without.
Kyllä	Ei	Kyllä

Predictors of in-hospital length of stay among cardiac patients: A machine learning approach	StratBAM:A Discrete-Event Simulation Model to Support Strategic Hospital Bed Capacity Decisions	Navigating hospitals safely through the COVID-19 epidemic tide: Predicting case load for adjusting bed capacity
Tahani A. Daghistani, Radwa Elshawi, Sherif Sakr, Amjad M. Ahmed, Abdullah Al-Thwayee, Mouaz H. Al-Mallah	Priyantha Devapriya & Christopher T.B.Strömblad & Matthew D.Bailey & Seth Frazier & John Bulger & Sharon T.Kemberling & Kenneth E Wood	Donker, T., Bürkin, F. M., Wolkewitz, M., Haverkamp, C., Christoffel, D., Kappert, O., Hammer, T., Busch, H. J., Biever, P., Kalbhenn, J., Bürkle, H., Kern, W. V., Wenz, F., & Grundmann, H.
2019	2015	2021
16917	23136	
96	12	2
machine learning	The Strategic Bed Analysis Model (StratBAM) = discrete-event simulation model	Static incidence model, Dynamic incidence model, Care pathway model
Random Forest (RF), Artificial Neural Network (ANN), Support Vector Machine (SVM) and Bayesian Network (BN).		
patients demographics (age and gender), vital sign on admission (heart rate, systolic and diastolic blood pressure, and body surface area), cardiovascular risk factors (hypertension, diabetes, dyslipidemia, smoking, and obesity), labs on admission (serum creatinine, high density lipoprotein, and ejection fraction), clinical diagnosis (congestive heart failure, acute coronary syndrome, acute myocardial infarction), and admission criteria (eligibility, seasonal admission, and physician experience)).	Arrival rates, Proportion of transfers, distribution of the length of care (LOC), distribution of the discharge times, Number of beds in each level of care unit, Bed Turnover time, The Number of beds, The volume and category of patients that can be transferred and cared for at other hospitals, Forecasts of patient volume and average LOS, Seasonality trends in patient arrivals, Patient wait time, Number of patients boarding, Occupancy rate	historialliset ja dataan perustuvat keskiarvot ja arviot tартtuvuudesta ja leviämisenopeudesta
The Random Forest (RF) model significantly outperformed all other models for the Accuracy (80%) and ROC 0.94	only minor differences between the model output overall average LOS (4.48 days) and the FY(fiscal year) 2013 actual LOS (4.44 days), a difference that is less than 1 h. Comparing the number of visits at each level of care as a percent of all visits, the difference between the operational data and the model output is within 0.5 %. The difference in discharges was within 0.6 %.	The static model, deployed before the epidemic, exaggerated the bed occupancy for general wards (116 forecasted vs 66 observed), ICUs (47 forecasted vs 34 observed), and predicted the peak too late: general ward forecast April 9 and observed April 8 and ICU forecast April 19 and observed April 8. After April 5, the dynamic model could be run daily, and its precision improved with increasing availability of empirical local data.
Kyllä	Kyllä	Kyllä

Forecasting Daily Patient Outflow From a Ward Having No Real-Time Clinical Data	A deep attention model to forecast the Length Of Stay and the in-hospital mortality right on admission from ICD codes and demographic data	Predictive Modeling for Geriatric Hip Fracture Patients: Early Surgery and Delirium Have the Largest Influence on Length of Stay
Gopakumar, S., Tran, T., Luo, W., Phung, D., & Venkatesh, S.	Gaspard Harerimana, Jong Wook Kim, Beakcheol Jang	Hecht G, Slee CA, Goodell PB, Taylor SL, Wolinsky PR.
2016	2021	2019
		A retrospective analysis
12141	47796	
60		24
machine learning	machine learning	A multivariate analysis, unadjusted regression analysis, linear regression analysis
(1) autoregressive integrated moving average (ARIMA), (2) the autoregressive moving average with exogenous variables (ARMAX), (3) k-nearest neighbor regression (kNN), (4) random forest regression (RF), and (5) support vector regression (SVR)	SMOTE (Synthetic Minority Oversampling Technique) used for unbalanced test data, Hierarchical Attention Network (HAN) using Natural language processing (NLP)	
Patient data: patient ID, age, gender, admission ID, name of the ward, entry and exit time, bed ID, patient class (21 categories), admission type (7 categories) Ward data: seasonality, trend, admissions, discharges, occupancy, admission type, patient referral, patient class, age category, number of wards visited, elapsed length of stay (eLoS)	Age, marital status, number of previous surgeries, smoking (y/n), admission type (ICD-9), time to previous discharge, admission location, previous LoS	age, gender, time to surgery, ASA score, CCI score, INR on admission, admission to the orthopaedic unit (e.g. cohorting patients successfully), and occurrence of delirium during the hospitalization.
In our experiments, forecast based on RF model outperformed all other models. Forecasting error rate is 31.9% for next day discharge (as measured by sMAPE) which is in the same ballpark as the recent work of [24], though we had no real-time clinical information.	Lyhyen ennakkoinnin (alle 10 päivää) tarkkuus: 90.3%, keskipitkän ennakkoinnin (10-30 päivää) tarkkuus: 92.5%, pitkän ennakkoinnin (yli 30 päivää) tarkkuus: 88.7% käyttäen HAN + SMOTE	the model made reasonable predictions for 63% our patients.
Kyllä	Kyllä	Vaatii lisätutkimusta

Predicting Length of Stay Across Hospital Departments	Short and Long term predictions of Hospital emergency department attendances	Predicting Length of Stay for Obstetric Patients via Electronic Medical Records
J. M. P. Gutiérrez, M. -Á. Sicilia, S. Sanchez-Alonso and E. García-Barriocanal	Tahseen Jilani, Gemma Housley, Graziela Figueredo, Pui-Shan Tang, Jim Hatton, Dominick Shaw	Gao C, Kho AN, Ivory C, Osmundson S, Malin BA, Chen Y.
2021	2019	2017
63932		9188
60	60	36
machine learning	Fuzzy Time Series (FTS)	machine learning
Parhaat tulokset: Decision Tree algorithm, Random Forest algorithm		Random forest (RF)
Gender, admission date, specialty of admission, discharge disposition, surgery (y/n), primary diagnosis, primary procedures, age, cost weight, groups related with diagnosis, (historic LoS for training data)	Historical attendance data. To handle weekday variations, the data was first segmented into each weekday time series and a separate model for each weekday was performed. Seasonality testing was performed, followed by Box-Cox transformations.	1) demographics (e.g., age), 2) encounter information (e.g., admission and discharge date), 3) diagnosis (e.g., billing codes) assigned to an encounter, and 4) clinical (e.g., obstetrics), ICD-9 codes
Tutkimus jakoi tulokset osastoihin, Decision tree ja Random Forest algoritmit olivat tarkimpia 64% ja 61% tarkkuudella. Molempien keskimääräinen ohiarviointi hoitojakson pituudesta oli alle päivän. (Soveltuvimmilla osastoilla, muilla osastoilla suurempia eroja)	Emergency department: forecasting daily admissions (in the next 4 weeks), the mean absolute percentage error ranged from 2.63% to 4.72%. Monthly presentations 2.09% to 2.81%	random forest regression model predicted LOS with an accuracy of 49% under an error range of 12 hours. Model based on the top 10% of ICD-9 codes can achieve an accuracy as high as those based on all involved ICD-9 codes (over 1800)
Vaatii lisätutkimusta	Kyllä	Vaatii lisätutkimusta



A predictive model for increased hospital length of stay following geriatric hip fracture	Recursive neural networks in hospital bed occupancy forecasting	Short-term forecasting of hospital discharge volume based on time series analysis
Olivia M. Knoll, Nikita Lakomkin, Michelle S. Shen, Moses Adebayo, Parth Kothari, Ashley C. Dodd, Basem Attum, Nathan Lee, Deepak Chonah, Manish K. Sethi	Ekaterina Kutafina, Istvan Bechtold, Klaus Kabino & Stephan M. Jonas	L. Luo, X. Xu, J. Li and W. Shen
2019	2019	2017
retrospective study		
614	353520	19696
108	161	36
negative binomial regression model (NBM)	nonlinear autoregressive model with exogenous terms (NARX)	machine learning
		random forests (RF)
age, race, sex,(ASA) classification, cigarette use, and alcohol use. Pre-operative comorbidities: diabetes, obesity, hypertension, myocardial infarction (MI), congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), liver disease, disseminated cancer, paralysis, peripheral vascular disease (PVD), pulmonary circulation disorders (PCD), thyroid disease, significant weight loss in the 6 months preceding surgery, renal failure, surgical intervention and admitting service (medicine or orthopaedics).	patient identifier, time of admission and discharge, and the name of the clinic the patient was admitted to. No personal information on the patients or staff was provided	demographic information (e.g., gender, age) and the date of discharge for each patient. From the data, we could easily sum up the number of discharged patients per day.
Hypertension (IRR: 1.10, p = 0.029) and disseminated cancer (IRR: 1.24, p = 0.007) were found to be significantly associated with LOS. In addition, two demographic/presenting variables, admission to the medicine service (IRR: 1.48, p < 0.001) and male sex (IRR: 1.09, p = 0.034), were shown to be independent risk factors for prolonged LOS. These variables were synthesized into a LOS formula, which estimated LOS to within 3 days of the true length of stay for 0.758 of the series (95% confidence interval: 0.661 to 0.855).	average MAPE on our testing set is 4.01, 5.55 and 5.48% on 1 day, 30 days and 60 days forecast respectively	The findings indicate that RF, the only multivariate model among the three models, performs best with a MAPE of 0.2095, and the results can be used to aid in strategic decision-making on inpatient beds resource planning in response to predictable discharges.
Vaatii lisätutkimusta	Kyllä	Vaatii lisätutkimusta

Predictive model of length of stay in hospital among older patients	Predicting Length of Stay in Intensive Care Units after Cardiac Surgery: Comparison of Artificial Neural Networks and Adaptive Neuro-fuzzy System	Predicting Emergency Department Inpatient Admissions to Improve Same-day Patient Flow
Radcliffe Lisk, Mahir Uddin, Anita Parbhoo, Keefai Yeong, David Fluck, Pankaj Sharma, Michael E. J. Lean & Thang S. Han	Hamidreza Maharlou, Sharareh R. Niakan Kalhori, Shahrbanoo Shahbazi, Ramin Ravangard	Jordan S. Peck MS, James C. Benneyan PhD, Deborah J. Nightingale PhD, Stephan A. Gaehde MD, MPH,
2019	2018	2012
	cross-sectional, analytical, and applied study	
374		5801
18		6
Statistical analysis	machine learning	Statistical analysis
Stepwise logistic regression (univariate & multivariate)	adaptive neuro-fuzzy inference system (ANFIS)	logit-linear regression model
sex, and pre-existing major co-morbidities including ischaemic heart disease (IHD), atrial fibrillation, stroke, Parkinson's disease, chronic obstructive pulmonary disease (COPD), malignancies, inflammatory bowel disease, diabetes mellitus and hypertension.	sex, LVEF, Previous cardiac operation, renal disease, cardiac hypertension, TIA or stroke, pulmonary disease, peripheral vascular disease, MIDCAB, embolism, smoking, hypercholesterolemia, COPD, Creatinine $\geq$ 2.0mg/Dl, Atrial fibrillation, preoperative infection, surgery type (CABG, MVR, AVR, other), CPR, type of operation, reoperation/re-exploration, NYHA, sinus rhythm, OPCAP, cardiogenic shock, heart valve surgery, mild valve pathology, age, aortic clamp time, BMI, hematocrit, duration of CPB (cardiopulmonary bypass)	patient age, primary complaint, bed type designation, and arrival mode
The adjusted risk of prolonged LOS was increased by threefold in those who were discharged to places other than usual residence, 10-fold in those admitted with acute stroke, sevenfold with first hip fracture and 10-fold with recurrent hip fractures, fourfold with acute infections, 10-fold with other acute fractures and 15-fold with malignancies (Table 1). Acute coronary syndrome (ACS) or COPD, diabetes, inflammatory bowel disease and or falls did not predict prolonged LOS. [total explained variance in LOS = 32.2%]	The results showed that the adaptive neuro-fuzzy algorithm (with mean squared error [MSE] = 7 and R = 0.88) resulted in the creation of a more precise model (LoS in days) than the artificial neural network (with MSE = 21 and R = 0.60).	area under the receiver operating characteristic curve of 0.887, an R2 of 0.58, and a daily average estimation error for the summative model of 0.19 beds
Vaatii lisätutkimusta	Kyllä	Vaatii lisätutkimusta

Real-time prediction of mortality, readmission, and length of stay using electronic health record data	Simulating Bed Capacity: Evaluating the Impact of Healthcare Service Transfers	Random Survival Forests for Predicting the Bed Occupancy in the Intensive Care Unit
Xiongcai Cai, Oscar Perez-Concha, Enrico Coiera, Fernando Martin-Sanchez, Richard Day, David Roffe, Blanca Gallego	R. Bares, J. Griffiths, V. Knight, J. Williams, K. Baboolal and A. Nelson	Ruysinck, J., van der Herten, J., Houthoof, R., Ongena, F., Couckuyt, I., Gadeyne, B., Colpaert, K., Decruyenaere, J., De Turck, F., & Dhaene, T
2015	2012	2016
32634		14480
42	12	57
Bayesian Network model		machine learning
	Discrete Event Simulation (DES)	Random Survival Forests (RSF)
Age, cumulative LoS prev. year, no. of days since previous admission, no. of days since admission, day since last event, pathology tests from admission, gender, mental disorder, cancer, triage code, day of week, time of day, marital code, mode of arrival to emergency, ward	arrival rates by hour of the day and day of the week, which are supplemented by monthly effects, patient group, gender, and age band	Sequential Organ Failure Assessment (SOFA) score = indicators of the coagulation, renal function, cardiovascular system function, respiratory function, liver function, and central nervous system function. (and the amount of days the patient has been admitted). 19 predictor variables
Predicting the next 7 days. The highest predictive power was achieved on day 1 (24 hours after a given prediction time), with a daily average accuracy of 86% and AUROC = 0.83, and decreased slightly with time. Daily average AUROC remained above 0.80 for all days. As with previous models, prediction of death was the most accurate outcome, with average accuracies of 93% and AUROC = 0.84.	average LOS was estimated as 4.77 days for elective patients and 10.72 days for emergency patients. This resulted in a total estimated requirement of approximately 61 beds at 85% occupancy, a percentage error of less than 2%. Neurosurgery bed capacity deterministically calculated for 85% occupancy achieved the 90th percentile of simulated bed usage, on average. This bed capacity did not meet demand on a day-to-day basis. Our software suggested that up to 12 additional beds were required to consistently meet demand at the 90th percentile.	Absolute mean error (AME) <1.5 when predicting LoS between 3 and 16 days
Kyllä	Vaatii lisätutkimusta	Kyllä

On the estimation of hospital beds occupancy after hip surgery	Non-Linear Feature Selection for Prediction of Hospital Length of Stay	Predictive model of length of stay and discharge destination in neuroscience admissions
S. D. Sousa, C. S. Rodrigues and E. P. Nunes	S. Kabir and L. Farrokhvar	Stecker, M. M., Stecker, M., & Falotico, J
2016	2019	2017
		retrospective correlational descriptive study
2030	715143	170
12	12	12
Statistical analysis	machine learning	Statistical analysis
logistic regression	non-linear feature selection, artificial neural network (ANN), support vector machines (SVM), and logistic regression (LR)	Bonferroni correction, analysis of variance (ANOVA), Spearman rank correlation, forward stepwise linear regression
Age, gender, ASA classification, Surgical Apgar Score - measured after surgery (predicts risk of complication), type of hip surgery, weekday of the surgery, starting hour of the surgery, duration of the surgery (in hours) - measured after	Katso liite 3. lähde: 10.1109/ICMLA.2019.00162	age, sex, height, weight, location prior to admission, principle diagnosis, various physiological measurements, comorbidity, independence in various ADLs prior to admission, LOS, and disposition upon discharge. where patient was admitted (stroke, neurointensive care, etc.), modified Rankin scale (mRS) on admission and discharge, tracheostomy or feeding tube during the stay, number of consults and if included palliative care, citizenship status, experienced a nosocomial infection, deep vein thrombosis.
The third and final model used all the available variables before and after surgery, including the patient Apgar and duration of the surgery. The third model resulted in a correct classification of 72.1% of cases. (Los > 7 days OR LoS < 7 days)	Katso liite 4. lähde: 10.1109/ICMLA.2019.00162 tarkkuus keskimäärin 80-91% parhaissa olosuhteissa	Kyseinen tutkimus ei arvioi hoitajakson pituutta tai ennakkoinnin tarkkuutta, vaan ehdottaa tärkeimpiä parametreja niiden arviointiin: Functional status at admission, specifically, a higher modified Rankin score and a lower systolic blood pressure along with the acquisition of deep vein thrombosis, catheter associated urinary tract infections, intubation, and admission to an intensive care unit all have a statistically significant effect on the hospital length of stay.
Vaatii lisätutkimusta	Kyllä	

A Model for Inpatient Ward Simulation and Forecasting	Predicting the length of stay at admission for emergency general surgery patients a cohort study
T. Latruwe, M. V. d. Wee, J. Devriese, P. Vanleenhove, S. Verbrugge and D. Colle	T.L. Ward, S.J. Raybould, A. Mondal, J. Lambert, B. Patel
2021	2021
	cohort study
	104
36	
	Multivariate analysis
Ward's method, CHI-square test	least squares regression, student's t-test, ANOVA
3M Service Line, APR-DRG, Severity of Illness (SOI), and average Length of Stay (aLoS), bed type	Age, Sex, Smoking status, English speaking, Early Warning Score (EWS), Individual Vital signs, Blood glucose, ASA score, Haemoglobin, White Cell Count, Platelets, C-reactive protein, Serum Creatinine, eGFR, Housing status (1 = Stable housing >6months, 2 = stable housing <6months, 3 = Supported housing, 4 = Unstable housing), Psychiatric medication use (1 = None, 2 = Antidepressants, 3 = Other psychiatric medication), Level of social support (1 = Living with partner, 2 = Sees close friend daily, 3 = Sees close friend < daily, 4 = No support), Barthel index (0-100)
CHI-square tests are used to compare the simulated inpatient frequency distribution with the observed distribution. In 62.3% percent of the 77 cases, the simulated distribution does not significantly differ from the observed distribution.	Study aim was to identify characteristics which predict LoS. Model achieved Root mean square error of 0.817041 and R <sup>2</sup> of 0.338829 and the the most important variables were: Barthel Index, Hb, Gender, Social support, Psych history, respiratory rate, housing status, age, english speaker, eGFR. The R2 showed that this model accounts for 34% of the variance in LOS in the cohort which is to be expected with a data set taken at admission.
Vaatii lisätutkimusta	Vaatii lisätutkimusta

### Liite 3: Kabir, S. & Farrokhvar L. tutkimuksen parametrit

#	Feature	Description	#	Feature	Description
1	AGE	Age	21	MORBPROB	Estimated Probability of Morbidity
2	ANESTHES	Principal anesthesia technique	22	MORTPROB	Estimated Probability of Mortality
3	ASACLAS	ASA classification	23	NSUPINFEC	Number of Superficial Incisional SSI Occurrences
4	ASCITES	Ascites	24	OPTIME	Total operation time
5	BLEEDDIS	Bleeding disorders	25	PRSEPIIS	Systemic Sepsis
6	DIABETES	Diabetes mellitus with insulin	26	RACE	New race
7	DIALYSIS	Currently on dialysis (pre-op)	27	RENAFAIL	Acute renal failure (pre-op)
8	DISCANCR	Disseminated cancer	28	SEX	Gender
9	DISCHDEST	Discharge destination	29	SMOKE	Current smoker within one year
10	DYSPNEA	Dyspnea	30	SSSIPATOS	Superficial Incisional SSI PATOS
11	ELECTSURG	Elective surgery	31	STEROID	Steroid use for chronic condition
12	EMERGNCY	Emergency case	32	SUPINFEC	Occurrences Superficial surgical site infection
13	ETHNICITY	Ethnicity Hispanic	33	SURGSPEC	Surgical Specialty
14	FNSTATUS2	Functional health status Prior to Surgery	34	TRANSFUS	Preop transfusion of in 72 hours prior to surgery
15	HEIGHT	Height in inches	35	TRANST	Transfer status
16	HtoODay	Days from Hospital Admission to Operation	36	VENTILAT	Ventilator dependent
17	HXCHF	Congestive heart failure (CHF) in 30 days before surgery	37	WEIGHT	Weight in lbs
18	HXCOPD	History of severe COPD	38	WNDCLAS	Wound classification
19	HYPERMED	Hypertension requiring medication	39	WNDINF	Open wound/wound infection
20	INOUT	Inpatient/outpatient	40	WTLOSS	>10% loss body weight in last 6 months

Kuva 1. Kabir, S. & Farrokhvar, L. tutkimuksessa mukana olleet parametrit.

Lähde: Kabir, S.; Farrokhvar, L. 2019. Non-Linear Feature Selection for Prediction of Hospital Length of Stay. Verkkoaineisto.

<<https://doi.org/10.1109/ICMLA.2019.00162>>. Luettu 14.9.2021.

## Liite 4: Kabir, S. &amp; Farrokhvar L. tutkimuksen tarkkuudet

Patient's Category	Classifier	All Features	Top 30	Top 25	Top 20	Top 15	Top 10	Top 5
General Surgery	ANN	0.8645	<b>0.9076</b>	0.8770	0.8538	0.8161	0.7652	0.7052
	SVM	0.8624	0.8461	0.8356	0.8352	0.8106	0.7445	0.6555
	LR	0.8402	0.8253	0.8184	0.7880	0.7522	0.7101	0.6089
Vascular	ANN	0.8848	0.8758	0.8815	0.8662	<b>0.9116</b>	0.7914	0.6712
	SVM	0.8833	0.8712	0.8415	0.8310	0.8005	0.7689	0.6356
	LR	0.8523	0.8225	0.8095	0.7891	0.7668	0.7217	0.6123
Urology	ANN	0.8032	0.8142	0.8367	<b>0.8563</b>	0.7750	0.7546	0.6615
	SVM	0.8024	0.8001	0.7923	0.7905	0.7624	0.7325	0.6112
	LR	0.7827	0.7523	0.7127	0.6913	0.6537	0.6158	0.5810
Plastics	ANN	0.8261	0.8467	0.8507	<b>0.8714</b>	0.7954	0.7681	0.7213
	SVM	0.8233	0.8211	0.8105	0.8070	0.7565	0.7165	0.6865
	LR	0.8040	0.7773	0.7528	0.7193	0.7018	0.6894	0.6405
Otolaryngology	ANN	0.7441	0.7623	<b>0.7988</b>	0.7572	0.7427	0.7243	0.6819
	SVM	0.7955	0.7609	0.7896	0.7493	0.7366	0.7083	0.6631
	LR	0.7348	0.7227	0.7066	0.6824	0.6439	0.6325	0.6071
Orthopedics	ANN	0.8361	<b>0.8634</b>	0.8229	0.7909	0.7656	0.7308	0.7006
	SVM	0.8342	0.8317	0.8138	0.7642	0.7467	0.7249	0.6407
	LR	0.8122	0.8019	0.7713	0.7506	0.7446	0.7106	0.6281
Gynecology	ANN	0.8775	<b>0.8921</b>	0.8623	0.8367	0.7945	0.7712	0.7569
	SVM	0.8635	0.8583	0.8313	0.8193	0.7914	0.7530	0.7014
	LR	0.8242	0.8234	0.8202	0.8242	0.8242	0.8238	0.8238
Neurosurgery	ANN	0.8530	0.8515	<b>0.8871</b>	0.8341	0.8121	0.7580	0.7351
	SVM	0.8418	0.8391	0.8111	0.8057	0.7762	0.7336	0.6982
	LR	0.8172	0.7960	0.7802	0.7575	0.7276	0.6862	0.6512
Other	ANN	<b>0.7973</b>	0.7818	0.7644	0.7681	0.7326	0.7014	0.6813
	SVM	0.7890	0.7637	0.7599	0.7655	0.7115	0.6965	0.6134
	LR	0.7625	0.7461	0.7079	0.6942	0.6698	0.6342	0.5911
All Patients	ANN	<b>0.8911</b>	0.8601	0.8366	0.8148	0.8031	0.7859	0.6718
	SVM	0.8691	0.8420	0.8189	0.8023	0.7815	0.7564	0.6477
	LR	0.8441	0.8163	0.7756	0.7432	0.7274	0.7011	0.6221

Kuva 1. Kabir, S. & Farrokhvar, L. tutkimuksen mallien tarkkuudet eriteltynä osastokohtaisesti. Lähde: Kabir, S.; Farrokhvar, L. 2019. Non-Linear Feature Selection for Prediction of Hospital Length of Stay. Verkkoaineisto. <<https://doi.org/10.1109/ICMLA.2019.00162>>. Luettu 14.9.2021.